ROBUST OPTIMIZATION FOR RENEWABLE ENERGY INTEGRATION IN POWER SYSTEM OPERATIONS

A Thesis
Presented to
The Academic Faculty

by

Álvaro Lorca

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Industrial and Systems Engineering

Georgia Institute of Technology
August 2016

ROBUST OPTIMIZATION FOR RENEWABLE ENERGY INTEGRATION IN POWER SYSTEM OPERATIONS

Approved by:

Professor Xu Andy Sun, Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Shabbir Ahmed
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Antonio J. Conejo
Department of Integrated Systems Engineering
Department of Electrical and Computer Engineering
The Ohio State University

Professor Arkadi Nemirovski
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Alexander Shapiro
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: 18 July 2016

To Alejandra

ACKNOWLEDGEMENTS

I would first like to express my most sincere gratitude to my advisor Dr. Xu Andy Sun. Working with Dr. Sun has been a true privilege. His continuous encouragement and patience have been fundamental for my development as an independent researcher. "Deep thinking" is not something that happens without effort, and this is only one of the many things I have been able to understand better from working with him. I feel very fortunate to have had such a great guidance throughout my doctoral studies.

I would also like to deeply thank the members of my Thesis committee for their service: Drs. Shabbir Ahmed, Antonio Conejo, Arkadi Nemirovski and Alexander Shapiro. I am truly lucky to have had such world-leading experts in different fields exposed to my research.

I am also grateful for the support from and collaboration with ISO New England, in particular, Drs. Jinye Zhao, Tongxin Zheng and Eugene Litvinov. I learned many things about power systems from them that would have been almost impossible to learn on my own. The internship program they have set up provides a wonderful experience; in particular, it was very meaningful for me to physically see bulk electric machinery with my own eyes and talk to power system operators.

I am also grateful for the support from PSERC. In particular, it was quite a learning experience to participate in the progress/proposal meetings.

I would also like to thank Drs. Pinar Keskinocak and Ozlem Ergun, with whom I have worked in the area of post-disaster logistics. My experience with them has also significantly helped to make of me a better researcher.

I am also grateful for the work of Drs. Gary Parker, Paul Kvam and Alan Erera as Chairs of Graduate Studies during my Ph.D., and to Pam Morrison and Amanda

Ford for their valuable administrative assistance. More than one long discussion with Dr. Parker helped me guide my way through the Ph.D.; his experience truly makes him a one-of-a-kind mentor.

I must also say that I am indebted to all of the friends me and my wife made over these wonderful years in Atlanta. The list is too long for names here! You guys know who you are, Chilean and non-Chilean. So much support in so many ways back and forth. I made friendships that I never thought were even possible.

Finally, and above everything, I would like to say that half of my Ph.D. is well deserved by my wife Alejandra. I am extremely fortunate to have found someone like her to enjoy the journey that life is. And to think that we are lucky enough to have come back to Chile with a beautiful three-month-old daughter, Fernanda. I have no words to describe the joy she is in our lifes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Optimization provides critical support for the operation of electric power systems. As power systems evolve, enhanced operational methodologies are required, and innovative optimization models have the potential to support them. The need for sustainability has led to many transformations, including the deep adoption of wind and solar energy in many power systems. These renewable energy sources have tremendous environmental benefits and can be very convenient economically, however, the power supply they provide is highly uncertain and difficult to predict accurately. This Thesis proposes Robust Optimization models and algorithms for improving the management of uncertainty in electric power system operations. The main goal is to devise new operational methodologies to support the integration of variable renewable energy sources.

The first part of this Thesis presents the development of an adaptive robust optimization model for the economic dispatch problem under uncertainty in wind power. The goal of this problem is to determine the power output levels of generating units in order to minimize costs while satisfying several technical constraints. The concept of dynamic uncertainty set is developed to account for temporal and spatial correlations in wind speeds. Further, a simulation platform is implemented to combine the dispatch model with statistical prediction tools in a rolling-horizon framework. Extensive numerical experiments are carried out on this platform using real wind data, showing the potential benefits of the proposed approach in terms of cost and reliability improvements over deterministic models and simpler robust optimization models that ignore temporal and spatial correlations.

The second part proposes a multistage adaptive robust optimization model for the unit commitment problem, under uncertainty in nodal net loads. The purpose of this problem is to schedule available generating capacities in each hour of the next day, including on/off generator decisions. The proposed model takes into account the time causality of the hourly unfolding of uncertainty in the power system operation process, which is shown to be relevant when ramping capacities are limited and net loads present significant variability. To deal with large-scale systems, the idea of simplified affine policies is explored and a solution method based on constraint generation is developed. Extensive computational experiments on a 118-bus test case and a real-world power system with 2736 buses demonstrate that the proposed algorithm is effective in handling large-scale power systems and that the proposed multistage robust model can significantly outperform a traditional deterministic model and an existing two-stage robust model in both operational cost and system reliability.

The third part develops a more sophisticated multistage robust unit commitment model, where the temporal and spatial correlations of wind and solar power are considered, as well as energy storage devices. A new specialized simplified affine policy is proposed for dispatch decisions, and an efficient algorithmic framework using a combination of constraint generation and duality based reformulation with various improvements is developed. Extensive computational experiments show that the proposed method can efficiently solve the problem on a 2736-bus system under high dimensional uncertainty of 60 wind farms and 30 solar farms. The computational results also suggest that the proposed model leads to significant benefits in both costs and reliability over robust models with traditional uncertainty sets as well as deterministic models with reserve rules.

Finally, the fourth part explores how to jointly consider the non-convexity of the power flow equations and the uncertainty in renewable outputs in power dispatch problems. Here, a two-stage adaptive robust optimization model is developed for the

alternating current optimal power flow problem, considering multiple time periods and including technical details such as transmission line capacities and the reactive capability curves of conventional generators and renewable units. To solve this challenging problem, it is proposed to use convex relaxations and an alternating direction method to identify worst-case uncertainty realizations. Further, a speed-up technique based on screening transmission line constraints is explored. Extensive computational experiments show that the solution method is efficient and that there are significant advantages both from the economic and reliability standpoints as compared to a deterministic model for this problem.

CHAPTER I

INTRODUCTION

Optimization plays a key role in the management of electric power systems. From real-time to long-term planning, the most critical power system decisions are supported by a variety of different optimization problems. In dealing with the uncertainty that affects these decisions, today's power system operators typically enhance deterministic optimization models through well-tuned ad-hoc rules that aim to maintain the reliability of the system and minimize costs. While this approach is effective, we can expect that the development of new techniques in the area of optimization under uncertainty could yield significant benefits to this practice. This is the challenge that guides this Thesis. More specifically, this Thesis proposes models and algorithms to address key optimization problems in electric power system operations by systematically addressing uncertainty through Robust Optimization.

The remainder of the Introduction will first provide motivation for the problems studied, and then outline an overview of this Thesis.

1.1 Motivation of this Thesis

The need for sustainability has led to significant transformations in the electric power industry. In the year 2000, the global installed wind power capacity was 17.4 GW, while in 2015 this capacity was 432.4 GW [43]. Solar power followed a similar trend, having the global installed photovoltaic capacity move from 1.4 GW to 237.3 GW in the same range of years [93]. These two variable renewable energy sources already represent a large proportion of the electricity generated in several power systems around the world, and they will continue to grow. For example, California's power mix in 2015 presented 9.4% wind power and 9.0% solar power [20], and renewable

portfolio standards will push these numbers higher in the future.

Unfortunately, the negligible variable costs and the tremendous environmental advantages of wind and solar power are coupled with a significant intermittency and unpredictability in their power output. Given this, the integration of wind and solar power requires careful operational decisions to ensure reliability under higher levels of uncertainty. A key aspect to this integration is the flexibility a power system presents. If there is enough controllable generation capacity that can quickly adapt to changes in renewable power output, the system can maintain energy balance effectively, while if this is not possible, renewable power might need to be curtailed for security reasons. Thus, a key question in integrating intermittent renewables is how to determine and employ available flexible generation capacities when they are limited or costly.

In the operation of power systems, the day-ahead unit commitment (UC) problem and the real-time economic dispatch (ED) problem are two key procedures for the management of wind and solar power. UC and ED are difficult problems that need to take into account the technical characteristics of generating units as well as the transmission grid. The goal of UC is to obtain a schedule of available generating capacities for each hour of the next day, which includes the determination of on/off decisions for thermal generating units. If more units are selected to be available, there will be more flexibility to withstand potential strong variations in wind and solar power, however, higher fixed costs will be required. This presents a challenging tradeoff to power system operators. Further, not only day-ahead decisions need to be taken carefully under wind and solar power. Every 5 minutes, or similar, power systems solve the ED problem, which consists of determining the power output levels of all available generators. Usually, wind and solar power can be much more accurately predicted in the short term, however, dispatch levels at different times are related through the capacity of generating units to ramp their output up or down, so ED decisions must also be carefully selected to withstand uncertainty in future renewable

power outputs. Depending on the dispatch levels of generators, these can provide more or less flexibility. For example, if a generator is operating at its maximum power output, this generator does not offer the flexibility to generate more power if renewable power drops down unexpectedly. This presents yet another critical tradeoff to the operation, as some cheap generators might need to be kept generating less power than its maximum output to prepare for renewable power variations.

In order to have UC and ED prepare for uncertainty in renewable power, and also in power demand, traditional deterministic models typically consider forecasts for these parameters and are enhanced through the concept of reserves. The idea is to optimize operational decisions assuming these parameters take their forecast values, but subject to maintaining additional generation capacities prepared in case more power than expected is needed. Many methods can be used for determining reserves, but uncertainty is not explicitly modeled in this approach. This raises the question of whether a more systematic method can be devised. Fortunately, the field of optimization under uncertainty offers different possibilities to optimize decisions having a good understanding of their consequences under parameters that are not known exactly.

There are several areas of optimization under uncertainty, among which Robust Optimization provides an effective paradigm under large-dimensionality of uncertainty, as is needed in power systems. In Robust Optimization uncertain parameters are modeled through uncertainty sets. The idea is to find decisions that maintain feasibility under any realization of the uncertain parameters in the uncertainty set, and such that the cost under a worst-case uncertainty realization is minimized. A very interesting and contingent recent development in Robust Optimization involves the concept of adaptive recourse decisions, that is, decisions that can change as required depending on the specific realization of uncertain parameters. This is very important because it provides a better understanding of how the flexibility of some

decisions can affect a given system. For example, in daily power system operations the on/off generator decisions have to be selected in the day-ahead (these can be seen as non-adaptive decisions) while power output decisions can be selected as wind and solar power outputs are revealed throughout the day (these can be seen as adaptive decisions). These are aspects that will be thoroughly discussed throughout this Thesis.

1.2 Overview of this Thesis

In this Thesis we develop concepts in the area of optimization under uncertainty to support the integration of intermittent renewable energy into power systems. The main goal is to develop Robust Optimization techniques to better manage the uncertainty in the power output of wind and solar farms and in power demand. More specifically, there are three fundamental research questions addressed:

- How to capture temporal and spatial correlations of wind and solar power in a robust optimization model for power dispatch decisions?
- How to develop a multistage adaptive robust optimization approach for power scheduling in a computationally tractable way?
- How to jointly consider the non-convexity of the power flow equations and the uncertainty in renewable outputs in power dispatch problems?

Chapter 2 explores the first question. This Chapter presents the development of an adaptive robust optimization model for the ED problem under uncertainty in wind power, considering multiple time periods. In this Chapter, the concept of dynamic uncertainty set is developed to account for temporal and spatial correlations in wind speeds. Further, a simulation platform is implemented to combine the ED model with statistical prediction tools in a rolling-horizon framework. Extensive numerical experiments are carried out on this platform using real wind data, showing the potential

benefits of the proposed approach in terms of cost and reliability improvements over deterministic models and simpler robust optimization models that ignore temporal and spatial correlations.

The second question is explored in Chapter 3, where a multistage adaptive robust optimization model is developed for the UC problem, under uncertainty in nodal net loads. The proposed model takes into account the time causality of the hourly unfolding of uncertainty in the power system operation process, which is shown to be relevant when ramping capacities are limited and net loads present significant variability. To deal with large-scale systems, the idea of simplified affine policies is explored and a solution method based on constraint generation is developed. Extensive computational experiments on a 118-bus test case and a real-world power system with 2736 buses demonstrate that the proposed algorithm is effective in handling large-scale power systems and that the proposed multistage robust UC model can significantly outperform the deterministic UC model and an existing two-stage robust UC model in both operational cost and system reliability.

Chapter 4 takes the first two questions further, through the development of a more sophisticated multistage robust UC model where the temporal and spatial correlations of wind and solar power are considered, as well as energy storage devices. A new specialized simplified affine policy is proposed for dispatch decisions, and an efficient algorithmic framework using a combination of constraint generation and duality based reformulation with various improvements is developed. Extensive computational experiments show that the proposed method can efficiently solve multistage robust UC problems on a 2736-bus system under high dimensional uncertainty of 60 wind farms and 30 solar farms. The computational results also suggest that the proposed model leads to significant benefits in both costs and reliability over robust models with traditional uncertainty sets as well as deterministic models with reserve rules.

Finally, Chapter 5 explores the third question. In this Chapter, a two-stage adaptive robust optimization model is developed for the alternating current (AC) optimal power flow problem (OPF), considering multiple time periods and including technical details such as transmission line capacities and the reactive capability curves of conventional generators and renewable units. To solve this challenging problem, it is proposed to use convex relaxations and an alternating direction method to identify worst-case uncertainty realizations. Further, a speed-up technique based on screening transmission line constraints is explored. Extensive computational experiments show that the solution method is efficient and that there are significant advantages both from the economic and reliability standpoints as compared to a deterministic model for AC OPF.

CHAPTER II

ADAPTIVE ROBUST OPTIMIZATION WITH DYNAMIC UNCERTAINTY SETS FOR MULTI-PERIOD ECONOMIC DISPATCH UNDER SIGNIFICANT WIND

2.1 Introduction

The exceptional benefits of wind power as an environmentally responsible energy resource have led to the rapid increase of wind energy in power systems all over the world. At the same time, wind energy possesses some characteristics drastically different from conventional generating resources in terms of high stochasticity and intermittency in production output. Due to this, deep penetration of wind power will introduce significant uncertainty to the short-term and real-time operation of power systems, in particular, to the day-ahead unit commitment (UC) and the real-time economic dispatch (ED) procedures. If the uncertainty of such variable resources is not managed properly, the system operator may have to face severe operating conditions such as insufficient ramping capabilities from the conventional generating resources due to the sudden strong loss of wind power, complicated by other contingencies, load surge, and transmission congestions [31]. These arising challenges call for new methods and models for power systems operation, and have attracted significant interests from both the electricity industry and academia.

The current UC and ED procedures rely on a combination of optimization tools and operational rules. The main optimization models used for UC and ED are deterministic models, where the uncertainties, such as demand, are assumed to take nominal forecast values. To deal with unexpected contingencies and sudden demand surge, the deterministic optimization model is complemented by operational rules

that require extra generation resources, the so-called reserves, to stay available for quick response. The discrepancy between the forecast and realization of uncertainty has been relatively small in power systems composed of conventional load and supply. However, as observed in the recent experience, operating power systems with high penetration of variable resources, especially wind power, requires new methods to deal with uncertainty. See [110] for an overview of the challenges of integrating wind in power systems from the perspective of UC, ED, frequency regulation and planning.

Facing these challenges, both industry and academia have devoted much effort to improving the current ED practice. In particular, dynamic dispatch models with look-ahead capabilities have gained renewed interests. The basic ideas can be traced back to [4] and [87]. Recent works have made significant advancement. [112] presents a look-ahead ED model with new statistical methods for wind forecast. The Mid-continent ISO has proposed look-ahead ED models with ramping products [76]. And [78] studies the selection of spinning reserve requirements under generation outages and forecast errors of demand and wind power. All these models can be characterized as deterministic ED models. Their simple optimization structure, improved performance, and closeness to the current operation make them appealing candidates to impact industry practice. This motivates the present Chapter to propose further advances and compare with these promising models.

Stochastic optimization has been a popular approach and extensively studied in the literature especially for the day-ahead unit commitment operation. For example, [97] proposes one of the first stochastic UC models. [108, 101] propose security constrained UC models and consider stochastic wind power generation. [19] presents a short-term forward electricity market-clearing model under net load uncertainty, for the purpose of allowing high penetrations of wind power while keeping the system secure. [74] deals with the selection of spinning and nonspinning reserves through a market-clearing model under stochastic wind power generation. [98] presents a

stochastic UC model for significant wind and shows the benefits of more frequent planning and over a deterministic approach. [81] studies reserve requirements for wind integration using a stochastic UC model. [82] proposes multiarea stochastic UC models for high wind penetration. [102] proposes a chance-constrained two-stage stochastic UC for systems with wind power uncertainty.

Regarding stochastic ED, the literature is much less extensive. [56] presents a stochastic programming model without recourse actions for a single-period ED problem. [64] presents a stochastic model of a single-period ED problem under post-contingency frequency constraints. [114] presents a chance-constrained look-ahead ED model where the probability of incurring lost load is constrained and a sampling based scenario approximation approach is used for dealing with wind power randomness, however, transmission constraints are not considered in this work to ease computational burden. We would like to note that UC and ED have quite significant differences in decision structures and therefore modeling considerations: the UC has a relatively clear two-stage decision making structure, whereas for ED, the modeling choices are more diverse. Constructing a stochastic ED model with proper decision structure and desirable computational properties merits further research efforts. The literature in this respect still leaves much room for new contributions.

Recently, robust optimization has emerged as an alternative methodology for optimization under uncertainty [6, 9]. Robust optimization provides several features that are particularly appealing to applications in power systems. In particular, the robust optimization approach seeks to optimize system performance in a controlled manner against the worst-case scenario, which is indeed consistent with the philosophy of the current operational practice; robust optimization provides a data-driven way to model uncertainty, which scales well with the increasing dimension of data and is flexible and practical for many situations; robust optimization models are usually computationally tractable for large-scale systems.

Recent works have proposed robust optimization models for UC problems [94, 53, 13, 118, 116]. [94] provides a robust formulation for the contingency constrained UC problem. [53], [13], [118] present two-stage adaptive robust models, with commitment decisions in the first stage and dispatch decisions in the second stage. In [13] a two-stage robust UC model with security constraints is formulated and tested on the power system operated by ISO New England. [53] deals with a formulation including pumped storage hydro under wind power output uncertainty. Hybrid models and alternative objectives have also been explored to mitigate the conservativeness of the robust solution [115, 54]. Efficient solution methods for the two-stage robust UC have been proposed [53, 13, 118, 63]. Recently, [63] presents acceleration techniques based on cutting planes and column generation for solving the two-stage robust UC problem under full transmission line constraints.

On the other hand, the benefit of robust optimization for the ED operation has not been fully explored. [119] presents a two-stage robust ED model for a single-period regulation dispatch problem, where the first stage corresponds to dispatch and regulation capacity decisions, and the second stage corresponds to the dispatch of automatic generation control (AGC), after observing demand. [51] recently proposes a robust optimal power flow model using affine policies for the AGC dispatch under renewable energy uncertainty. Affine policy is an approximation to the fully adaptive policy used in [119]; however, as argued in [51], affine dependence on uncertainty may be a more suitable form for AGC dispatch. The work in [111] applies two advanced statistical methods for wind forecasting, and integrates these models with a robust look-ahead ED. However, their model is of a static robust nature, which lacks the adaptability of a two-stage robust model proposed here; their model also relies on the existing types of uncertainty sets, which will be significantly improved by a new type of uncertainty sets proposed in this Chapter.

If we try to summarize the above works, we can draw the following observations:

1) there is a great amount of interests to improve the ED practice; in particular, the recently developed look-ahead ED models have attracted considerable attention in both academia and industry; 2) the existing works on power system operation under uncertainty have focused on UC problems in a day-ahead operating environment, while both stochastic and robust ED models are relatively less explored; 3) the existing robust UC and ED models have used a similar type of uncertainty sets, which we call *static* uncertainty sets, whereas it is important to start considering uncertainty sets that can capture the highly dynamical and correlated variable resources such as wind power.

In this Chapter, we propose new robust optimization models for system dispatch under high wind penetration. In particular, the contributions of this Chapter are summarized below:

1. We propose a two-stage adaptive robust optimization model for the multi-period ED, which has a different decision structure from the existing two-stage robust UC and robust ED models. The proposed robust ED model is designed for a rolling-horizon operational framework to model the real time ED process.

2. We introduce a new type of uncertainty sets, the *dynamic uncertainty sets*, as a modeling technique to account for the dynamic relationship between uncertainties across decision stages. Such uncertainty sets explicitly model temporal and spatial correlations in variable sources. We also propose a data-driven approach to construct such dynamic uncertainty sets, which is simple to implement in practice.

3. We develop a comprehensive simulation platform, which integrates the proposed robust ED model with statistical procedures for constructing dynamic uncertainty sets using real-time data. Extensive experiments are performed on this platform.

The remainder of this Chapter is organized as follows. Section 2.2 introduces dynamic uncertainty sets and discusses practical construction methods. Section 2.3 proposes the adaptive robust multi-period ED model and solution methods. Section 2.4 presents the simulation platform and the evaluation framework. Section 4.5 shows extensive computational experiments to demonstrate the effectiveness of our approach. Finally, Section 4.6 concludes.

2.2 Dynamic uncertainty sets

In robust optimization, uncertainty is modeled through uncertainty sets, which are the building blocks of a robust optimization model and have direct impact on its performance. We may summarize three criteria for constructing uncertainty sets as follows. A well constructed uncertainty set should 1) capture the most significant aspects of the underlying uncertainty, 2) balance robustness and conservativeness of the robust solution, and 3) be computationally tractable.

2.2.1 Static uncertainty sets

Previous works on robust UC have focused on static uncertainty sets, and have treated uncertainty resources of different characteristics in an aggregated, indistinguishing way, see for example [13, 53, 118]. More specifically, consider the following uncertainty set for net demand vector $\boldsymbol{d}_t = (d_{1t}, ..., d_{N^d t})$:

$$\mathcal{D}_t = \Bigg\{ \boldsymbol{d}_t : \sum_{j \in \mathcal{N}^d} \frac{|d_{jt} - \overline{d}_{jt}|}{\hat{d}_{jt}} \leq \Gamma^d \sqrt{N^d}, \\ d_{jt} \in [\overline{d}_{jt} - \Gamma^d \hat{d}_{jt}, \overline{d}_{jt} + \Gamma^d \hat{d}_{jt}] \; \forall \, j \in \mathcal{N}^d \Bigg\}, \tag{1}$$

where $\mathcal{N}^d, N^d$ denote the set and the number of loads, and d_{jt} is the net demand of load j at time t. According to (1), d_{jt} lies in an interval centered around the nominal value $\overline{d}_{jt}$ with a width determined by the deviation $\hat{d}_{jt}$. Further, the size of the uncertainty set is controlled by Γ^d. If $\Gamma^d = 0$, $\mathcal{D}^t = \{\overline{\boldsymbol{d}}_t\}$, corresponding to

a singleton set of the nominal demand. As Γ^d increases, more demand vectors are contained in the uncertainty set, thus increasing the protection of the robust solution against larger demand variations.

The above uncertainty set is called *static* uncertainty set, because the uncertainties at later time periods are *independent* of those in earlier periods. That is, the dynamics of uncertainty evolution over time is not explicitly captured. Some recent work proposed additional budget constraints over time periods (e.g. [53, 118]). The modified uncertainty set imposes a coupling of uncertainty between time periods and uncertain sources, however, similar to (1), it still does not directly characterize the temporal and spatial correlations of uncertainty; also, by coupling through the entire horizon, the realization of uncertainty breaks the time causality with past depending on the future realization. Yet another drawback of existing models is that uncertain sources of different nature are treated indistinguishably. For example, the uncertainty characteristics of wind power output are different from those of the conventional load, yet the existing proposals consider aggregated net load as the primitive uncertainty [13, 53, 118]. Demand uncertainty is usually much less pronounced and less dynamic than wind, therefore, a static uncertainty set as (1) is an appropriate model. However, it is important to explore well suited uncertainty models for wind, especially for high level penetration of such variable resources.

2.2.2 Dynamic uncertainty sets

To explicitly model the correlation between multiple uncertain resources within one time period as well as the dynamics of each uncertain resource evolving over time periods, we propose the following general form of uncertainty sets, called *dynamic uncertainty sets*: For each time t,

$$\Xi_t(\boldsymbol{\xi}_{[1:t-1]}) = \left\{ \boldsymbol{\xi}_t : \exists \boldsymbol{u}_{[t]} \text{ s.t. } f(\boldsymbol{\xi}_{[t]}, \boldsymbol{u}_{[t]}) \leq \mathbf{0} \right\}, \tag{2}$$

where $\boldsymbol{\xi}_{[t_1:t_2]} \triangleq (\boldsymbol{\xi}_{t_1}, \ldots, \boldsymbol{\xi}_{t_2})$ and in shorthand $\boldsymbol{\xi}_{[t]} \triangleq \boldsymbol{\xi}_{[1:t]}$. In (2), the uncertainty vector $\boldsymbol{\xi}_t$ explicitly depends on uncertainty at stages before time t and the $\boldsymbol{u}$'s are auxiliary variables, $f(\boldsymbol{\xi}_{[t]}, \boldsymbol{u}_{[t]})$ is a vector of convex functions that characterize the dynamics of uncertainty evolution. For the uncertainty set to be computationally tractable, f should be semi-definite representable [6].

As an illustrative example, the dynamic uncertainty set could represent a dynamic interval for $\boldsymbol{\xi}_t$:

$$\boldsymbol{\xi}_t \in \left[\underline{\boldsymbol{\xi}}_t(\boldsymbol{\xi}_{[t-1]}), \overline{\boldsymbol{\xi}}_t(\boldsymbol{\xi}_{[t-1]})\right],$$

where the upper and lower bounds of the interval at time t, namely $\underline{\boldsymbol{\xi}}_t(\boldsymbol{\xi}_{[t-1]})$ and $\overline{\boldsymbol{\xi}}_t(\boldsymbol{\xi}_{[t-1]})$, are functions of uncertainty realizations in previous time periods, rather than fixed values as in static uncertainty sets (1).

A simple and useful specialization of (2) is the linear dynamic uncertainty set, given as

$$\sum_{\tau=1}^{t} (\boldsymbol{A}_\tau \boldsymbol{\xi}_\tau + \boldsymbol{B}_\tau \boldsymbol{u}_\tau) \leq \boldsymbol{0}, \tag{3}$$

which mimics linear dynamics and is also computationally tractable. In the following, we will propose a specific method for constructing linear dynamic uncertainty sets using time series analysis tools.

2.2.3 Constructing dynamic uncertainty sets for wind power

The proposed dynamic uncertainty set (2) is very general. In this section, we present a specific method to construct a dynamic uncertainty set for wind power using linear systems (3). The key idea is to fuse time series models with the concept of dynamic uncertainty sets.

We denote the wind speed vector of multiple wind farms at time t as $\boldsymbol{r}_t = (r_{1t}, \ldots, r_{N^w t})$, where r_{it} is the wind speed at wind farm i and time t. Define the

dynamic uncertainty set for $\boldsymbol{r}_t$ as:

$$\mathcal{R}_t(\boldsymbol{r}_{[t-L:t-1]}) = \Big\{\boldsymbol{r}_t :\ \exists\, \widetilde{\boldsymbol{r}}_{[t-L:t]},\ \boldsymbol{u}_t \quad \text{s.t.}$$

$$\boldsymbol{r}_\tau = \boldsymbol{g}_\tau + \widetilde{\boldsymbol{r}}_\tau \quad \forall \tau = t-L, \ldots, t \tag{4a}$$

$$\widetilde{\boldsymbol{r}}_t = \sum_{s=1}^{L} \boldsymbol{A}_s \widetilde{\boldsymbol{r}}_{t-s} + \boldsymbol{B}\boldsymbol{u}_t \tag{4b}$$

$$\sum_{i \in \mathcal{N}^w} |u_{it}| \le \Gamma^w \sqrt{N^w} \tag{4c}$$

$$|u_{it}| \le \Gamma^w \qquad \forall i \in \mathcal{N}^w \tag{4d}$$

$$\boldsymbol{r}_t \ge \boldsymbol{0}\Big\}, \tag{4e}$$

where vectors $\boldsymbol{r}_{t-L}, \ldots, \boldsymbol{r}_{t-1}$ are the realizations of wind speeds in periods $t-L, \ldots, t-1$. Eq. (4a) decomposes wind speed vector $\boldsymbol{r}_\tau$ as the sum of a seasonal pattern $\boldsymbol{g}_\tau$, which is pre-estimated from wind data, and a residual component $\widetilde{\boldsymbol{r}}_\tau$ which is the deviation from $\boldsymbol{g}_\tau$. Eq. (4b) is the key equation that represents a linear dynamic relationship involving the residual $\widetilde{\boldsymbol{r}}_t$ at time t, residuals realized in earlier periods $t-L$ to $t-1$, and an error term $\boldsymbol{u}_t$. The parameter L sets the relevant time lags. In Eq. (4b), matrices $\boldsymbol{A}_s$'s capture the temporal correlation between $\boldsymbol{r}_t$ and $\boldsymbol{r}_{t-s}$, and $\boldsymbol{B}$ specifically captures the spatial relationship of wind speeds at adjacent wind farms at time t. Eq. (4c)-(4d) describe a budgeted uncertainty set for the error term $\boldsymbol{u}_t$, where Γ^w controls its size, and (4e) avoids negative wind speeds. $\mathcal{N}^w$ and N^w denote the set and number of wind farms, respectively.

Using the above uncertainty sets (4) for wind speeds, we can further construct dynamic uncertainty sets for wind power through power curve approximations. In particular, we denote the *available wind power* of wind farm i at time t as $\overline{p}^w_{it}$. Given the wind speed r_{it}, $\overline{p}^w_{it}$ is described by the following constraints

$$\overline{p}^w_{it} \ge h^0_{ik} + h_{ik}\, r_{it} \quad \forall i \in \mathcal{N}^g,\ k = 1, \ldots, K, \tag{5}$$

where parameters h^0_{ik}, h_{ik} are determined based on a convex piecewise linear approximation with K pieces of the increasing part of the power curve at wind farm i (in

our experiments, we use the power curve of GE 1.5MW wind turbine to approximate the aggregated output of a wind farm). Although (5) allows available wind power to exceed $\max_k\{h^0_{ik} + h_{ik}r_{it}\}$, the robust optimization model described in Section 2.3 will always ensure that the available wind power lies on the power curve including the plateau part for wind speed exceeding a cut-off value.

The dynamic uncertainty set of the available wind power $\overline{\boldsymbol{p}}^w_t$ is thus defined as

$$\overline{\mathcal{P}}^w_t(\boldsymbol{r}_{[t-L:t-1]}) = \Big\{\overline{\boldsymbol{p}}^w_t :\ \exists \boldsymbol{r}_t \in \mathcal{R}_t(\boldsymbol{r}_{[t-L:t-1]}) \text{ s.t. (5) is satisfied}\Big\}, \tag{6}$$

based on which we can define the uncertainty set for the trajectory of available wind power in time periods 2 through T, namely $\overline{\boldsymbol{p}}^w = (\overline{\boldsymbol{p}}^w_2, \ldots, \overline{\boldsymbol{p}}^w_T)$, as

$$\overline{\mathcal{P}}^w = \Big\{(\overline{\boldsymbol{p}}^w_2, \ldots, \overline{\boldsymbol{p}}^w_T) :\ \exists(\boldsymbol{r}_2, \ldots, \boldsymbol{r}_T) \text{ s.t. } \boldsymbol{r}_t \in \mathcal{R}_t(\boldsymbol{r}_{[t-L:t-1]}) \text{ and (5) is satisfied for } t = 2, \ldots, T\Big\}, \tag{7}$$

which is used in the robust ED model.

As a summary, we propose dynamic uncertainty sets (4) and (6) to capture the intrinsic temporal dynamics and spatial correlations of the wind power. We also distinguish wind power uncertainty from conventional demand uncertainty, which is modeled by traditional uncertainty sets (1). The proposed dynamic uncertainty set formulation (2) is quite general. The specific models for wind speed (4) and wind power (6) present one example for its implementation. Other models may be constructed using more sophisticated statistical tools. For example, the coefficient matrices $\boldsymbol{A}_s$ and $\boldsymbol{B}$ can be made time dependent as $\boldsymbol{A}_{st}$ and $\boldsymbol{B}_t$ using dynamic vector autoregression methods. Also, $\boldsymbol{r}_t$ can be replaced by a nonlinear transformation of wind speed to improve estimation accuracy. However, there is always a tradeoff between model complexity and performance. Our experiments show the above simple models (4)-(7) achieve a substantial improvement over existing static uncertainty sets.

See more discussion in Section 2.4.1 for parameter estimation and possible extensions for the dynamic uncertainty sets.

2.3 Adaptive robust economic dispatch formulation and solution method

2.3.1 Mathematical formulation

In this section, we propose an adaptive robust optimization model for the multi-period ED problem. In particular, the ED problem with T time periods is formulated as a two-stage adaptive robust model in the following way. The first-stage of the robust ED model comprises the current time period $t = 1$, while the second-stage comprises future time periods $t = 2, ..., T$. In the first-stage, the decision maker observes demand and available wind power at the current time period, and determines the dispatch solution, which will be implemented right away for time $t = 1$. Given the first-stage decision, the second-stage of the robust ED model computes the worst-case dispatch cost for the remaining time periods in the dispatch horizon. The overall robust ED model minimizes the total cost of dispatch at the current time period and the worst-case dispatch cost over the future periods.

We denote $\boldsymbol{x} = (\boldsymbol{p}_1^g, \boldsymbol{p}_1^w)$ as the vector of first-stage dispatch decisions, composed of generation of thermal units ($\boldsymbol{p}_1^g$) and wind farms ($\boldsymbol{p}_1^w$). Note that we allow wind generation to be dispatchable. The uncertainty includes both conventional load $\boldsymbol{d} = (\boldsymbol{d}_2, \ldots, \boldsymbol{d}_T) \in \mathcal{D}$ described by (1) and the available wind power $\overline{\boldsymbol{p}}^w = (\overline{\boldsymbol{p}}_2^w, \ldots, \overline{\boldsymbol{p}}_T^w) \in \overline{\mathcal{P}}^w$ described by the dynamic uncertainty set (7). The second-stage decisions are dispatch solutions $\boldsymbol{y} = (\boldsymbol{p}_t^g, \boldsymbol{p}_t^w, \forall t = 2, \ldots, T)$.

Mathematically, the two-stage robust multi-period ED model is formulated as follows,

$$\min_{\boldsymbol{x} \in \Omega_1^{det}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} + \max_{\boldsymbol{d} \in \mathcal{D}, \overline{\boldsymbol{p}}^w \in \overline{\mathcal{P}}^w} \min_{\boldsymbol{y} \in \Omega(\boldsymbol{x}, \boldsymbol{d}, \overline{\boldsymbol{p}}^w)} \boldsymbol{b}^\top \boldsymbol{y} \right\}, \tag{8}$$

where the first and second-stage costs are defined as

$$\boldsymbol{c}^\top \boldsymbol{x} = \sum_{i \in \mathcal{N}^g} C_i^g p_{i1}^g + \sum_{i \in \mathcal{N}^w} C_i^w p_{i1}^w$$
$$\boldsymbol{b}^\top \boldsymbol{y} = \sum_{t=2}^{T} \left(\sum_{i \in \mathcal{N}^g} C_i^g p_{it}^g + \sum_{i \in \mathcal{N}^w} C_i^w p_{it}^w \right),$$

where $\mathcal{N}^g$ denotes the set of generators, and C_i^g, C_i^w denote the variable costs of thermal generators and wind farms. We use linear dispatch costs, but it is straightforward to extend to piecewise linear approximations of nonlinear cost functions.

The feasible region Ω_1^{det} of the first-stage decision variables corresponds to the constraints of a single-period dispatch problem, that is

$$\Omega_1^{det} = \Bigg\{ \boldsymbol{x} = (\boldsymbol{p}_1^g, \boldsymbol{p}_1^w) : \ \underline{p}_{i1}^g \le p_{i1}^g \le \overline{p}_{i1}^g \quad \forall\, i \in \mathcal{N}^g \tag{9a}$$
$$0 \le p_{i1}^w \le p_i^{w,max} \quad \forall\, i \in \mathcal{N}^w \tag{9b}$$
$$p_{i1}^w \le \overline{p}_{i1}^{w,det} \quad \forall\, i \in \mathcal{N}^w \tag{9c}$$
$$-RD_i^g \le p_{i1}^g - p_{i0}^g \le RU_i^g \quad \forall\, i \in \mathcal{N}^g \tag{9d}$$
$$-RD_i^w \le p_{i1}^w - p_{i0}^w \le RU_i^w \quad \forall\, i \in \mathcal{N}^w \tag{9e}$$
$$\left| \boldsymbol{\alpha}_l^\top (\boldsymbol{E}^g \boldsymbol{p}_1^g + \boldsymbol{E}^w \boldsymbol{p}_1^w - \boldsymbol{E}^d \boldsymbol{d}_1^{det}) \right| \le f_l^{max} \ \forall\, l \in \mathcal{N}^l \tag{9f}$$
$$\sum_{i \in \mathcal{N}^g} p_{i1}^g + \sum_{i \in \mathcal{N}^w} p_{i1}^w = \sum_{j \in \mathcal{N}^d} d_{j1}^{det} \Bigg\}, \tag{9g}$$

where $\underline{p}_{it}^g, \overline{p}_{it}^g$ are the minimum and maximum power outputs of thermal generator i at time t; $p_i^{w,max}$ is the maximum power output at wind farm i, representing the cut-off level of the power curve; $\overline{p}_{i1}^{w,det}$ denotes the available wind power of wind farm i observed at current time $t = 1$; RD_i^g, RU_i^g are the ramp-down and ramp-up rates of thermal generators (similarly, RD_i^w, RU_i^w for wind farms); $\mathcal{N}^l$ is the set of transmission lines; $\boldsymbol{\alpha}_l$ is the network shift factor for line l; $\boldsymbol{E}^d, \boldsymbol{E}^g, \boldsymbol{E}^w$ are the network incidence matrices for loads, thermal generators and wind farms; f_l^{max} is the flow limit on line l; d_{j1}^{det} denotes the observed electricity demand at load j and time

$t = 1$. Constraints (9a), (9b) and (9c) enforce generation limits for thermal generators and wind farms, with (9c) ensuring that generation of wind farms does not exceed the available wind power at time $t = 1$. (9d) and (9e) enforce ramping rate limits for thermal generators and wind farms. (9f) represents line flow limits. (9g) represents energy balance.

Constraints in the second-stage problem are parameterized by the first-stage decision variables and uncertain parameters realized in the uncertainty sets. The feasible region of the second-stage dispatch decison $\boldsymbol{y} = (\boldsymbol{p}_t^g, \boldsymbol{p}_t^w, \forall t = 2, \ldots, T)$ is defined as

$$\Omega(\boldsymbol{x}, \boldsymbol{d}, \overline{\boldsymbol{p}}^w) = \Bigg\{ \boldsymbol{y} : \quad \text{s.t.} \ \ \forall t = 2, \ldots, T$$

$$\underline{p}_{it}^g \leq p_{it}^g \leq \overline{p}_{it}^g \quad \forall\, i \in \mathcal{N}^g, \tag{10a}$$

$$0 \leq p_{it}^w \leq p_i^{w,max} \quad \forall\, i \in \mathcal{N}^w, \tag{10b}$$

$$p_{it}^w \leq \overline{p}_{it}^w \quad \forall\, i \in \mathcal{N}^w, \tag{10c}$$

$$-RD_i^g \leq p_{it}^g - p_{i,t-1}^g \leq RU_i^g \quad \forall\, i \in \mathcal{N}^g, \tag{10d}$$

$$-RD_i^w \leq p_{it}^w - p_{i,t-1}^w \leq RU_i^w \quad \forall\, i \in \mathcal{N}^w, \tag{10e}$$

$$\left| \boldsymbol{\alpha}_l^\top (\boldsymbol{E}^g \boldsymbol{p}_t^g + \boldsymbol{E}^w \boldsymbol{p}_t^w - \boldsymbol{E}^d \boldsymbol{d}_t) \right| \leq f_l^{max} \ \forall\, l \in \mathcal{N}^l \tag{10f}$$

$$\sum_{i \in \mathcal{N}^g} p_{it}^g + \sum_{i \in \mathcal{N}^w} p_{it}^w = \sum_{j \in \mathcal{N}^d} d_{jt} \Bigg\}, \tag{10g}$$

where (10a)-(10g) are similar constraints as in (9), except that they are enforced for each time period $t = 2, \ldots, T$. Notice that (10b)-(10c) ensure that the dispatched wind generation is upper bounded by the minimum between the cut-off level $p_i^{w,max}$ and the available wind power $\overline{p}_{it}^w$. Also note that the first-stage dispatch decision is involved in constraints (10d)-(10e) to satisfy ramping constraints.

A few remarks are in order. First, (8) is a fully adaptive robust optimization model, namely the second-stage dispatch decision adapts to every realization of the uncertainty in the best possible way, which is similar to the existing robust UC model proposed in [13]. Second, there is a key difference between the two-stage structure

of the proposed robust ED (8) and the existing two-stage robust UC models. In particular, the decision stages of (8) correspond to the actual time periods, so that the first-stage decision can be directly used in the dispatch at the current period, and the dispatch decisions in the second stage can be re-optimized in the following periods. In comparison, the two-stage robust UC models have UC decisions in the first stage and dispatch decisions in the second stage, both for the *entire horizon*. Third, the two-stage structure of the robust ED model makes it convenient to incorporate into the real-time dispatch procedure. In particular, the robust ED model can be implemented in a rolling horizon framework; the dynamic uncertainty sets can also be updated periodically when new information is available. Fourth, the use of the DC power flow is consistent with the industry practice [33] and recent works in robust ED [119, 51]. AC power flow feasibility can be enforced by introducing an AC power flow module. Thus, to emphasize the key proposal of this Chapter, we keep the simple DC power flow model. Fifth, the robust ED model can also readily include convex piecewise linear costs.

2.3.2 Solution method

Several methods have been reported in the literature for solving two-stage adaptive robust optimization problems [13, 53, 113]. In [53], a Benders decomposition approach is proposed to solve the outer level problem and an exact method for the second-stage problem. In [113], a constraint and column generation (C&CG) technique is proposed and rigorously analyzed; an exact method using mixed-integer reformulations is proposed for the second-stage problem. In [13], a modified Benders decomposition framework is proposed for the outer level problem with an efficient heuristic method for the second-stage problem. The key modification to the traditional Benders decomposition is to add the worst-case extreme point and the associated dispatch constraints to the outer level problem in each iteration (see [13, Section IV]). This is similar to

the idea behind constraint and column generation in [113].

Problem (8) can be equivalently stated as:

$$\min_{\boldsymbol{x},\eta} \left\{ \boldsymbol{c}^\top \boldsymbol{x} + \eta : \ \eta \geq Q(\boldsymbol{x}), \ \boldsymbol{x} \in \Omega_1^{det} \right\}, \tag{11}$$

with

$$Q(\boldsymbol{x}) = \max_{\boldsymbol{\xi} \in \Xi} \ \min_{\{\boldsymbol{y}: \ \boldsymbol{G}\boldsymbol{y} \geq \boldsymbol{h} - \boldsymbol{E}\boldsymbol{x} - \boldsymbol{M}\boldsymbol{\xi}\}} \ \boldsymbol{b}^\top \boldsymbol{y}, \tag{12}$$

where $\boldsymbol{\xi} = (\boldsymbol{d}, \overline{\boldsymbol{p}}^w)$, $\Xi = \mathcal{D} \times \overline{\mathcal{P}}^w$, and the feasible region $\{\boldsymbol{y} : \boldsymbol{G}\boldsymbol{y} \geq \boldsymbol{h} - \boldsymbol{E}\boldsymbol{x} - \boldsymbol{M}\boldsymbol{\xi}\}$ represents the dispatch constraints in (10). Problem (11) is equivalent to:

$$\min_{\boldsymbol{x} \in \Omega_1^{det}, \eta, \{\boldsymbol{y}_l\}} \quad \boldsymbol{c}^\top \boldsymbol{x} + \eta \tag{13a}$$

$$\text{s.t.} \quad \eta \geq \boldsymbol{b}^\top \boldsymbol{y}_l \quad \forall \, l \tag{13b}$$

$$\boldsymbol{E}\boldsymbol{x} + \boldsymbol{G}\boldsymbol{y}_l \geq \boldsymbol{h} - \boldsymbol{M}\boldsymbol{\xi}_l^* \quad \forall \, l, \tag{13c}$$

where $\{\boldsymbol{\xi}_l^*\}_{l=1}^M$ is the set of extreme points of Ξ, and for each l, $\boldsymbol{y}_l$ is a vector of second-stage decisions associated to $\boldsymbol{\xi}_l^*$. (13) is the outer level problem, which shows a nice structure suitable for constraint generation. Indeed, (13) can be efficiently solved by adding $(\boldsymbol{\xi}_l^*, \boldsymbol{y}_l)$ and the associated constraints iteratively [13, 113].

In every iteration of this algorithm, $Q(\boldsymbol{x})$ must be evaluated, which involves solving a nonconvex max-min problem. Previous work has dealt with this problem using outer-approximation techniques [13] and exact methods based on mixed-integer programming (MIP) reformulations [53, 118, 113]. As will be demonstrated in the computational experiments (Section 2.5.5), the MIP method is time consuming for solving (12). Instead, we apply a simple "alternating direction algorithm" [60]. Taking the dual over the inner min in (12) we obtain

$$Q(\boldsymbol{x}) = \max_{\boldsymbol{\xi} \in \Xi, \boldsymbol{\pi} \in \Pi} \quad \boldsymbol{\pi}^\top (\boldsymbol{h} - \boldsymbol{E}\boldsymbol{x} - \boldsymbol{M}\boldsymbol{\xi}), \tag{14}$$

where $\Pi = \{\boldsymbol{\pi} \geq \boldsymbol{0} : \boldsymbol{\pi}^\top \boldsymbol{G} = \boldsymbol{b}\}$. For this *bilinear* program with separate polyhedral feasible regions Ξ and Π, the alternating direction algorithm optimizes over $\boldsymbol{\pi}$ with

$\boldsymbol{\xi}$ fixed, then over $\boldsymbol{\xi}$ with $\boldsymbol{\pi}$ fixed, and alternates; each of these iterations solves a linear program which achieves the optimum at an extreme point of the corresponding polyhedron Ξ or Π. The alternating algorithm is formally presented below.

Algorithm 1 Alternating Direction (AD) algorithm

Start with some $\boldsymbol{\xi}' \in \Xi$
repeat
  Solve $(*)$: $C \leftarrow \max_{\boldsymbol{\pi} \in \Pi} \ \boldsymbol{\pi}^\top (\boldsymbol{h} - \boldsymbol{E}\boldsymbol{x} - \boldsymbol{M}\boldsymbol{\xi}')$
  if $C < \infty$ **then**
    Let $\boldsymbol{\pi}'$ be an optimal solution of $(*)$
    Solve $C' \leftarrow \max_{\boldsymbol{\xi} \in \Xi} \ \boldsymbol{\pi}'^\top (\boldsymbol{h} - \boldsymbol{E}\boldsymbol{x} - \boldsymbol{M}\boldsymbol{\xi})$ and let $\boldsymbol{\xi}'$ be its optimal solution
  else
    $C' \leftarrow \infty$
  end if
until $C' = \infty$ **or** $C' - C \leq \delta$
output: C' as estimate of $Q(\boldsymbol{x})$ with solution $\boldsymbol{\xi}'$

This alternating direction method always converges to a KKT point of (14). The proof is omitted to save space. Section 2.5.5 also shows empirical evidence that this heuristic achieves good solution quality and fast convergence on the second-stage problem, comparing to the MIP method.

The overall two-level algorithm is presented in Fig. 1.

2.4 Simulation platform and evaluation metrics

In this Section, we describe the simulation platform and evaluation metrics for the proposed robust model. The motivation is to have a realistic simulation environment that integrates the dispatch optimization model with data analysis procedures which dynamically update the parameters in the optimization and uncertainty models. Fig. 2 illustrates the simulation process.

The simulation process is implemented in a rolling horizon framework. At each time period, the robust ED model is solved over a time window of T time periods. The first-stage dispatch solution for the current time period is implemented, while the second-stage dispatch solutions for remaining periods are not materialized; the

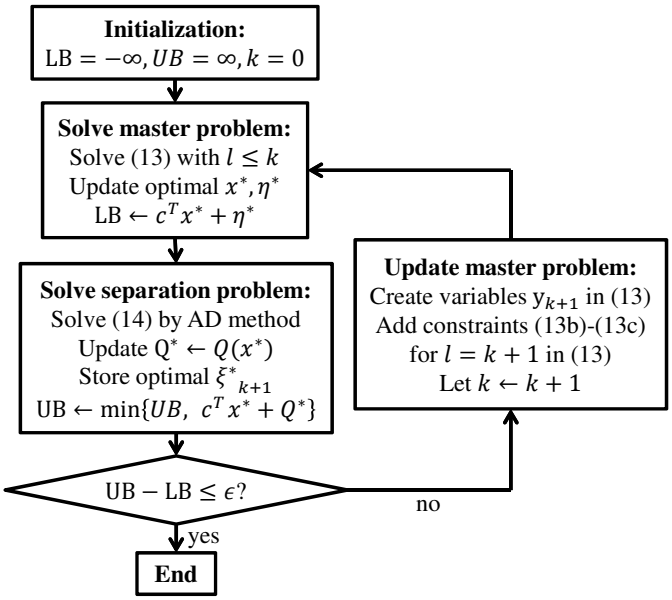

Figure 1: Flow chart for the overall two-level algorithm.

time horizon rolls forward by one time interval, where new realizations of demand and available wind power are observed, and dynamic uncertainty sets are periodically re-estimated and updated with the new observations (see Section 2.4.1). In order to focus the comparison on the ED policies, the simulation process uses a simplified UC schedule where all thermal generators are on all the time. In the future, we would like to extend the simulation framework to integrate UC decisions into the policy evaluation.

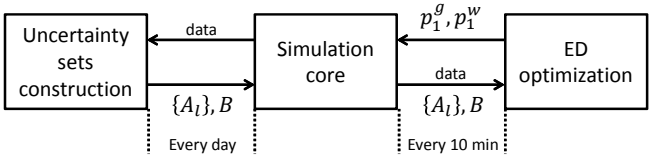

Figure 2: Simulation platform integrating ED optimization engine and data analysis tools for uncertainty model construction.

We compare different ED models by evaluating the average and standard deviation (std) of the production cost for every 10 minutes dispatch interval, which includes both generation cost and penalty cost resulting from the use of expensive fast-start units or load shedding.

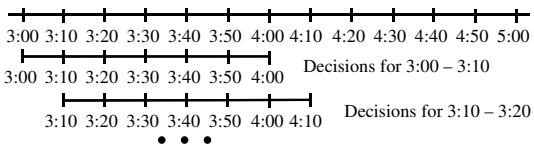

Figure 3: Concept of rolling horizon with 10 minute time periods and $T = 6$

2.4.1 Estimating the parameters of the dynamic uncertainty set for wind speeds

In order to estimate the parameters of model (4), consider the following time series model:

$$\boldsymbol{r}_t = \boldsymbol{g}_t + \widetilde{\boldsymbol{r}}_t \quad \forall t \tag{15a}$$

$$\widetilde{\boldsymbol{r}}_t = \sum_{s=1}^{L} \boldsymbol{A}_s \widetilde{\boldsymbol{r}}_{t-s} + \boldsymbol{\epsilon}_t \quad \forall t, \tag{15b}$$

where $\boldsymbol{r}_t$ is the vector of wind speeds at time t, $\boldsymbol{g}_t$ corresponds to a deterministic seasonal pattern, and $\widetilde{\boldsymbol{r}}_t$ corresponds to the deviation of $\boldsymbol{r}_t$ from $\boldsymbol{g}_t$. In this model, $\widetilde{\boldsymbol{r}}_t$ follows a multivariate autoregressive process of order L, determined by the innovation process $\{\boldsymbol{\epsilon}_t\}$, where $\boldsymbol{\epsilon}_t$ is a vector of normal random variables with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$, and vectors $\boldsymbol{\epsilon}_t$ are independent across different time periods.

Once seasonal patterns have been identified, parameter $\boldsymbol{g}_t$ can be determined. For example, daily and semi-daily seasonalities could be used. In such a case, using a 10 min time interval we would have $g_{it} = a_i + b_i \cos(\frac{2\pi t}{24\times 6}) + c_i \sin(\frac{2\pi t}{24\times 6}) + d_i \cos(\frac{2\pi t}{12\times 6}) + e_i \sin(\frac{2\pi t}{12\times 6})$ (since 24×6 is the number of time periods in a day). Parameters a_i, b_i, c_i, d_i, e_i can be estimated using linear regression [112].

The parameters of the autoregressive component $\widetilde{\boldsymbol{r}}$, namely the matrices $\boldsymbol{A}_s$ and $\boldsymbol{\Sigma}$, can be estimated using statistical inference techniques developed for time series [86], for which many computational packages are available. $\boldsymbol{B}$ in (4) is obtained from the Cholesky decomposition of $\boldsymbol{\Sigma}$.

The linear dynamic model (4) and the associated estimation method are appealing in their simple structure, which serves well our goal to demonstrate the concept of

dynamic uncertainty sets. Computational results also confirm their promising performance. Meanwhile, it is worth noting that the framework of dynamic uncertainty sets is flexible enough to incorporate more sophisticated statistical models, such as the ones proposed in [75], where autoregressive processes are fitted to nonlinearly transformed wind speeds. Using a piecewise linear approximation similar to the one proposed in (5), but this time for the transformed wind speed and wind power output, a dynamic uncertainty set can be again constructed using linear constraints.

2.5 Computational experiments

We conduct extensive computational experiments on the simulation platform to compare the proposed robust ED model and dynamic uncertainty sets with existing robust and deterministic dispatch models. The experiments are performed on the 14-bus and 118-bus IEEE test systems, both of which are modified to incorporate significant wind penetration. In the following, we introduce the detailed data for the 14-bus system, and present test results in Sections 2.5.1 to 2.5.5. The test results on the 118-bus system give a similar picture as the 14-bus system. The details are given in Section 2.5.6.

Table 1 summarizes P_{min}, P_{max}, 10-min ramping rates, and production costs of all three generators in the 14-bus system. The total generation capacity is 500MW. The system has 20 transmission lines and 11 conventional loads. The daily system demand is between 132.6MW and 319.1MW with an average of 252.5MW. The system has 4 wind farms, each with a capacity of 75MW (equivalent of 50 units of GE 1.5MW wind turbines). The total power output at each wind farm is approximated by a piecewise linear function of wind speed using the power curve data [37].

The wind speed data is obtained from [21] for four geographically adjacent locations with a 10-minute data interval. The average wind speeds at the four wind farms

Table 1: Thermal generators in 14-bus system

Gen	**Pmax (MW)**	**Pmin (MW)**	**Ramp (MW/10min)**	**Cost ($/MWh)**
1	300	50	5	20
2	100	10	10	40
3	100	10	15	60

are 4.8, 5.6, 5.1, 5.5 m/s, respectively. Using the power curve, the average total available wind power is 104.2 MW, equivalent to a 34.7% capacity factor, which is about 32.7% of peak demand and 20% of conventional generation capacity, representing a realistically high level of wind penetration. After removing stationary components, wind speeds at different sites present strong auto and cross correlation at several lags, which implies that the temporal and spatial dependencies are significant.

The proposed robust ED model has 9 time periods with a 10-min interval for each period (i.e. 1.5-hour look ahead). The robust ED model is evaluated on the simulation platform in the rolling-horizon framework. In particular, it is solved every 10 minutes over 35 days, for which real wind data is used for all wind farms. On each of the 35 days, the simulation engine updates the parameters of the dynamic uncertainty sets (4) using the available wind data up to that day. The penalty cost is $C^+ = 6000$ \$/MWh for under-generation, and $C^- = 600$ \$/MWh for over-generation [82, 115].

The simulation platform is implemented in a Python environment, interfaced with Cplex 12.5. Each robust ED takes less than a second to solve, and the entire simulation of 5040 periods takes about 40 minutes on a PC laptop with an Intel Core i3 at 2.1 GHz and 4GB memory.

Before presenting details, we first give a summary of the experiments and main results. We compare the proposed robust ED with dynamic uncertainty sets versus (1) deterministic look-ahead dispatch and its variant with reserve rules; (2) robust dispatch with static uncertainty sets. The experiments show that adaptive robust

ED with dynamic uncertainty sets significantly outperforms both alternative models by substantially reducing average production cost, the variability of the costs, and the probability of shortage events. Our experiments also show that the robust ED provides a *Pareto frontier* for the tradeoff between cost and reliability, which provides an informative guideline for choosing uncertainty set parameters and system operating points.

2.5.1 Robust ED versus look-ahead ED

In this section, we compare the proposed adaptive robust ED (Rob-ED) with the deterministic look-ahead dispatch (LA-ED). The robust ED model uses dynamic uncertainty sets (4) and (6) with 6 time lags i.e. $L = 6$. The parameter Γ^w controls the size of the uncertainty sets. Notice that when $\Gamma^w = 0$, the uncertainty set contains only one path of the forecasted wind speeds, the robust ED thus reduces to the LA-ED model.

2.5.1.1 Cost and reliability performance

Table 2 shows the performance of the two models: Column 2 for LA-ED, and Columns 3 to 7 for Rob-ED with different Γ^w's. The best average total cost of the Rob-ED model is achieved at $\Gamma^w = 0.5$, where the average cost of Rob-ED is 7.1% lower than that of LA-ED; at the same time, Rob-ED is able to reduce the standard deviation of the cost by 41.2%. We can also see that as Γ^w increases to 1.0, the robust ED can reduce the std of cost by 82.1%, with the average cost reduced by 3.75%. The shortage event frequency of the robust ED model is decreased by up to 80.1% and the associated penalty cost is reduced by 97.3% at $\Gamma^w = 1.0$. The change in penalty costs also implies that Rob-ED incurs less amount of constraint violation than LA-ED, when penalty occurs. The results show that the robust ED model is effective at improving economic efficiency and reducing risk associated with the dispatch solution, where the risk exactly comes from the highly uncertain wind power. As will be shown

in Section 2.5.6, more significant savings on cost and improvement over reliability are achieved for the 118-bus system.

Table 2: Performance of robust and deterministic ED

	LA-ED	Rob-ED				
Γ^w	**0.0**	0.1	0.3	**0.5**	0.7	**1.0**
Total Cost Avg ($)	**771.1**	758.5	734.0	**716.0**	718.2	**742.2**
Total Cost Std ($)	**1231**	1172	1000	**723**	513	**221**
Penalty Avg ($)	**88.2**	77.1	54.2	**30.6**	15.8	**2.4**
Penalty Freq (%)	**1.41**	1.21	0.95	**0.67**	0.46	**0.28**

2.5.1.2 Operational insights

We also want to gain some insights about the operational characteristics of the robust model. Table 3 shows average thermal generation (Therm avg) and wind generation (Wind avg) of the two models. We can see that the robust ED model on average tends to increase the use of thermal generation and curtail wind output: At $\Gamma^w = 0.5$, Therm avg is up by 4.3% and Wind avg down by 8.1%, comparing to LA-ED; at $\Gamma^w = 1.0$, Therm avg is up by 16.1% and Wind avg is down by 24.9%.

Fig. 4 shows a typical snapshot from simulation. Available wind power starts a fast and large drop at 21:30 (green curve), the deterministic LA-ED runs short of ramping capacity and incurs a spike of penalty cost (blue curve), while the system under robust ED is much less affected by this sudden wind event (red curve). The example shows that when the system has significant wind penetration, properly balancing wind and thermal generation becomes very important for system reliability.

The insight is the following. The two-stage robust ED computes wind scenarios over the future periods that are the most detrimental to the system, and makes the optimal dispatch solution to prepare the system against these scenarios. The worst-case wind scenarios often correspond to scenarios with large wind variation between periods as shown in Fig. 4. The robust ED model hedges against the potential large swing of wind by increasing thermal generation and moderately curtailing some wind

Table 3: Operational Aspects of Robust and Deterministic ED

	LA-ED	Rob-ED				
Γ^w	**0.0**	0.1	0.3	**0.5**	0.7	**1.0**
Therm avg (MW)	**164.6**	165.2	167.5	**171.7**	178.6	**191.1**
Wind avg (MW)	**87.9**	87.2	85.0	**80.8**	74.0	**61.5**

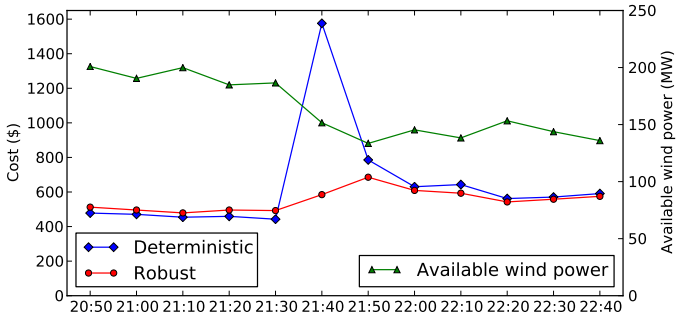

Figure 4: A snapshot of the product cost of LA-ED and Rob-ED with $\Gamma^w = 0.5$ when available wind power suddenly drops down.

output. In this way, the system maintains enough ramping capability to deal with potential sudden loss of available wind power. The balance between thermal and wind generation is controlled by the value of Γ^w of the uncertainty sets as shown in Table 3. In other words, the robust ED determines the optimal ramping schedule of thermal generators, rather than resorting to prefixed operation rules.

2.5.1.3 Comparing to look-ahead ED with reserve

Reserve is an engineering approach to handle net load uncertainty in a deterministic ED model. Typically, when UC is solved, reserve levels for the next day are co-optimized, and later in real time operation, reserves are used in cases of unexpected net load variations and other contingencies. Consider the following look-ahead ED model with reserve requirement (Res-LA-ED). The LA-ED model is complemented with reserve variables $R_{it} \in [0, \overline{R}_{it}]$, equations (9a) and (10a) are replaced by

$$\underline{p}^g_{it} \leq p^g_{it} \leq \overline{p}^g_{it} - R_{it} \quad \forall\, i \in \mathcal{N}^g,\, t = 1, \ldots, T,$$

Table 4: Performance of look-ahead ED with reserve

	LA-ED	Res-LA-ED			Rob-ED
ResFactor (%)	0	2.5	5	10	$\Gamma^w = 0.5$
Cost Avg ($)	771.1	770.0	773.3	790.3	716.0
Cost Std ($)	1231	1223.8	1211.8	1155.1	723
Penalty Avg ($)	88.2	86.7	84.8	71.6	30.6
Penalty Freq (%)	1.41	1.45	1.69	1.35	0.67

and the following reserve requirement constraints are added:

$$\sum_{i \in \mathcal{N}^g} R_{it} \geq R_t^{req} \quad \forall\, t = 1, \ldots, T.$$

We test the performance of this model for different reserve requirement levels R_t^{req}. We select R_t^{req} as a fraction of the total forecasted net load at time t (i.e. forecast of total demand minus total available wind power), and modify this proportion, denoted as "ResFactor" [82]. Table 4 presents the performance of Res-LA-ED under different values of ResFactor, as well as that of Rob-ED with $\Gamma^w = 0.5$.

From these results we can see that this reserve rule can improve the performance of LA-ED in both cost effectiveness and reliability, when the reserve requirement is properly chosen (ResFactor at 2.5%). As ResFactor increases, the reliability (Cost Std) keeps improving with the tradeoff of an increasing Avg Cost; the penalty cost and frequency are also reduced.

If we compare Res-LA-ED with Rob-ED, we can observe that the performance of Rob-ED is significantly better than the best of the three Res-LA-ED test cases: the Cost Avg is reduced by at least 7.14% (against ResFactor = 2.5%); the Cost Std is improved by at least 37.4% (against ResFactor = 10%); the penalty cost is reduced by at least 57.2%, and the penalty frequency is reduced by at least 50.3% (both against ResFactor = 10%).

2.5.2 Dynamic uncertainty sets versus static uncertainty sets

In this section, we compare adaptive robust ED equipped with dynamic uncertainty sets with the same robust ED model using static uncertainty sets. The goal is to study the benefits of dynamic uncertainty sets for modeling dynamic relations of wind power uncertainty across time stages and spatial locations.

We use dynamic uncertainty sets (4) with $L = 6$ as before (denoted as "DUS"), and construct two static uncertainty sets: one ignores the temporal correlation in (4) (denoted as "SUS1"), the other further ignores spatial correlations (denoted as "SUS2"). Note that both SUS1 and SUS2 are special cases of the dynamic uncertainty sets for $L = 0$, i.e. the uncertainty sets at different time intervals are independent of each other. To have a fair comparison, both in SUS1 and SUS2, $\boldsymbol{g}_t$ is improved after estimating $\boldsymbol{B}$ to force a persistent forecast of wind speeds for the nominal trajectory (improving the accuracy of the nominal trajectory considered).

Fig. 5 plots the standard deviation of the cost per 10 minutes interval (x-axis) versus the average of this cost (y-axis) for DUS, SUS1 and SUS2 with different values of Γ^w. On each curve, the right most point corresponds to $\Gamma^w = 0$, i.e. the deterministic LA-ED model. As Γ^w increases, both the average and std of the cost start to decrease, then after a certain apex value of Γ^w around 0.4 to 0.5, the std keeps decreasing but the average cost starts to increase. This behavior endows a "U" shape for all three curves. Every point on the right half of the "U" shape for Γ^w smaller than the apex value can be strictly improved in both average and std of cost by increasing Γ^w, while every point on the left half of the "U" shape cannot be strictly improved without trading off between average and std of the cost. In other words, on the right half of the curve, each point is dominated by the points to its left, whereas on the left half, no point is dominated by any other. Therefore, the left part of each curve shows the *Pareto frontier* of cost average vs cost standard deviation performance of the associated robust ED model. The system should be operated on the Pareto frontier.

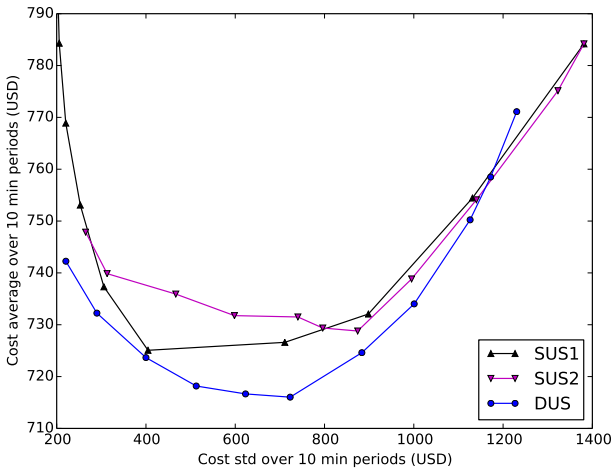

Figure 5: Cost std and cost average obtained for the policies determined by the different models with $\Gamma^w = 0.0, 0.1, ..., 1.0$

This provides an informative guideline for choosing a proper Γ^w.

Comparing the Pareto frontiers of the three uncertainty sets in Fig. 5, we can see that the dynamic uncertainty set has the lowest Pareto frontier, which means that to retain a same level of average cost, the robust ED with dynamic uncertainty sets achieves the lowest std (i.e. the highest reliability); or, to maintain a same level of std (i.e. reliability), the robust ED with dynamic uncertainty sets incurs the lowest cost. That is, robust ED with DUS *dominates* robust ED with static uncertainty sets. Between the two static uncertainty sets, SUS1 (that considers spatial correlation) dominates SUS2, which has neither temporal nor spatial correlation.

The static uncertainty set SUS2 is the first budgeted uncertainty set proposed in the literature [14] and has inspired its application in modeling net load uncertainty [13]. Works in [53, 118] further introduced budget constraints over time periods to limit the total variations of uncertain demand over the entire or part of the planning horizon. Now, we compare these static uncertainty sets with additional time budgets with DUS. It is worth emphasizing that the fundamental difference between DUS and SUS remains the same for DUS and SUS with time budgets.

We modify the uncertainty sets SUS1 and SUS2 with the following time budget

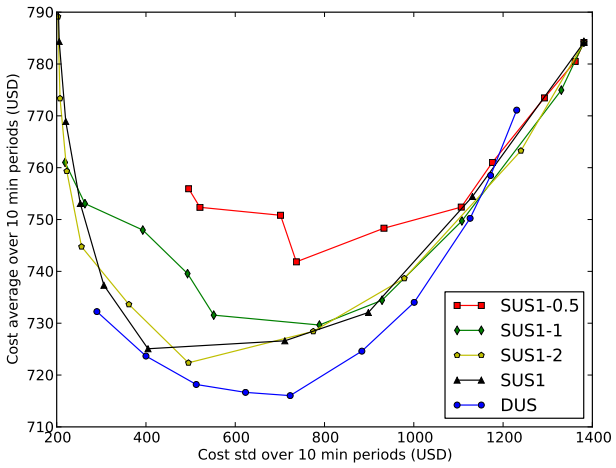

Figure 6: Cost std and cost average obtained for the policies determined by DUS and SUS1, with $\Gamma^T = 0.5, 1, 2$ for SUS1 and with $\Gamma^w = 0.0, 0.1, ..., 1.0$.

constraint:

$$\sum_{t=2}^{T} \sum_{i \in \mathcal{N}^w} |u_{it}| \leq \Gamma^T \Gamma^w \sqrt{N^w} \sqrt{T-1},$$

where $T = 9$ is the number of periods in the multi-period Rob-ED, and time budget parameter $\Gamma^T = 0.5, 1, 2$. Note that static uncertainty sets without time budget are equivalent to one with a large time budget as when $\Gamma^T \geq \sqrt{8}$ the time budget constraint becomes redundant.

Fig. 6 plots the std of cost per 10 min interval (x-axis) versus the average of that cost (y-axis) for DUS and SUS1 with additional time budgets. The curve denoted by SUS1-0.5 means the SUS1 uncertainty set with time budget $\Gamma^T = 0.5$ and Γ^w varies from 0.0 to 1.0. Among the three curves based on SUS1 with time budgets, we can see that Rob-ED achieves a better Pareto frontier for higher values of time budget (the red curve for SUS1-0.5 is dominated by the green curve for SUS1-1, which is further dominated by SUS1-2). SUS1 without time budget (or equivalently with a time budget $\Gamma^T \geq \sqrt{8}$) has a frontier comparable to the SUS1-2. Furthermore, all four SUS1 based curves are clearly dominated by the DUS curve.

Fig. 7 presents a similar comparison for SUS2 with time budgets. Here, the dominance of DUS over static uncertainty sets with time budgets is more eminent.

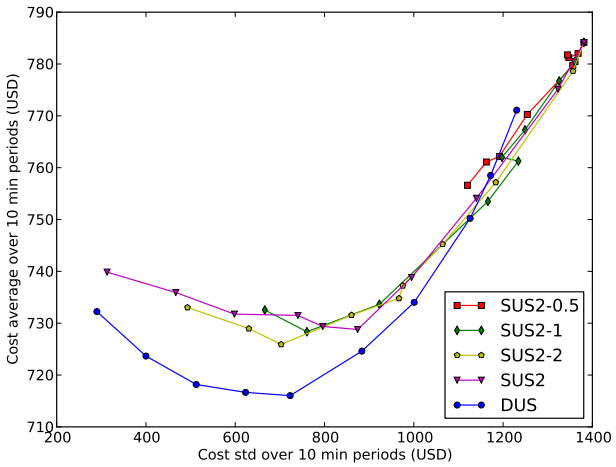

Figure 7: Cost std and cost average obtained for the policies determined by DUS and SUS2, with $\Gamma^T = 0.5, 1, 2$ for SUS2 and with $\Gamma^w = 0.0, 0.1, ..., 1.0$ for all policies

2.5.3 Impact of system ramping capacity

In this section, we study the relationship between system ramping capacity and the performance of robust ED models. The intuition is that higher ramping rates better prepare the system to deal with high variation of wind output. We want to see how much benefit the robust ED model provides under different system ramping capacities. Fig. 8 summarizes the computational results for three scenarios: base case with no change in ramping rates, and -25% or $+25\%$ change on each generator's ramping rates.

We can see that the robust ED model saves the average cost by 7.1% in the base case (the same numbers as in Section 2.5.1) comparing with the look-ahead ED; the saving increases to 21.2% for the reduced ramping case; even for the system with 25% more ramping for every generator, the robust ED still demonstrates a 3.7% saving in average cost over LA-ED. This demonstrates the clear benefit of Rob-ED over a wide range of system ramping conditions.

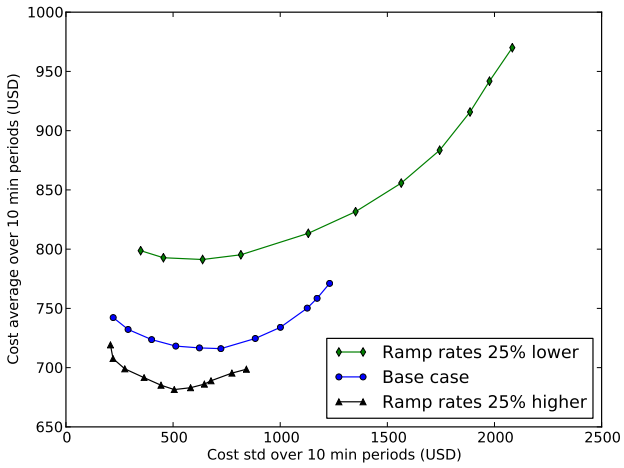

Figure 8: Cost std and cost average obtained for the Rob-ED with DUS for $\Gamma^w = 0.0, 0.1, ..., 1.0$, under modified ramping rates.

2.5.4 Considering both demand and wind uncertainty

In this section, we further incorporate traditional demand uncertainty into the robust ED model, using the static uncertainty sets (1), where $\overline{d}_{jt}$ and $\hat{d}_{jt}$ are selected as the mean and std of demand from previously realized values. The parameter Γ^d limits the total deviation of demand from its forecast. In simulation, the demand d_{jt} of each load j at each time period t is independently generated as a normal random variable with a std that equals a 5% of its mean, and is truncated to be nonnegative. Therefore, the generated random demand can be outside the uncertainty set. The choice of Γ^d controls the size of the demand uncertainty set.

Fig. 9 presents the performance of Rob-ED with dynamic uncertainty set for wind and static uncertainty sets (1) for load, at different values of Γ^d, Γ^w. At $\Gamma^d = 0$, the uncertainty set for demand is a singleton containing the forecast value, i.e. only wind uncertainty is considered (blue curve). By considering an uncertainty set for load with $\Gamma^d = 1$, the cost-reliability curve is shifted downward to the green curve, which consistently dominates the blue curve. The two curves are quite close though, which shows that wind is the dominating factor of uncertainty; the dynamic uncertainty set for wind significantly improves the system performance, while further incorporating

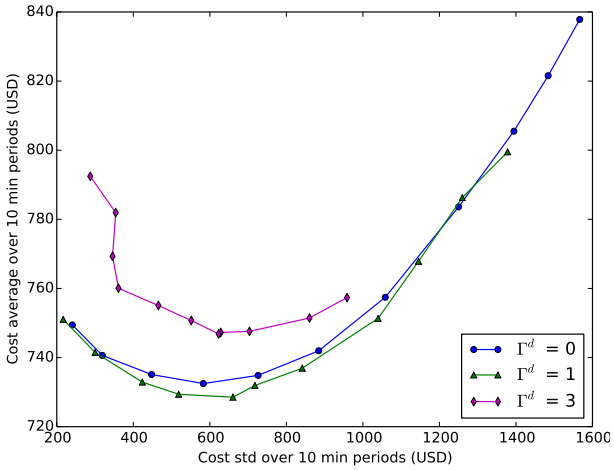

Figure 9: Cost std and average obtained by the Rob-ED with $\Gamma^w = 0.0, ..., 1.0$ and $\Gamma^d = 0, 1, 3$.

load uncertainty improves the performance modestly. The purple curve for $\Gamma^d = 3$ shows that too much conservatism in the load uncertainty model leads to inferior solutions. It again demonstrates that properly choosing the level of conservativeness of the uncertainty sets is critical to getting the best performance of the robust ED model. In particular, the best robust ED policy obtained by setting $\Gamma^d = 1, \Gamma^w = 0.6$ reduces the average cost by 13.1% lower than that of the deterministic LA-ED with $\Gamma^d = \Gamma^w = 0$, and reduces the std of the cost by 58.1%. This makes the robust ED model very attractive.

2.5.5 Performance of the alternating direction method for solving the second-stage problem

As discussed in Section 2.3.2, the proposed algorithm requires solving a bilinear program (14) in each iteration of the outer master problem. Therefore, to practically tackle large-scale problems, a fast and reliable method for the inner problem is needed. An alternating direction (AD) method is proposed in Section 2.3.2 for this purpose. In the literature, several exact MIP methods are proposed to solve the second-stage bilinear program (e.g. [53, 118, 113].) The MIP methods in [53, 118] rely on the special structure of the uncertainty sets used in their models, which are not shared

by the dynamic uncertainty sets. The exact MIP method proposed in [113] is based on the KKT conditions, which are applicable to general polyhedral uncertainty sets. Thus, we compare the AD algorithm to this MIP method.

In the experiment, we run the Rob-ED model in the rolling-horizon simulator for a 5-day horizon. This involves solving 720 Rob-ED models of the form (8), which amounts to 1529 inner bilinear programs (14). Every time, the bilinear program is solved by both the AD algorithm and the exact MIP method. We compare both running times and solution qualities.

The AD algorithm achieves convergence for all 1529 instances, and the average running time is 0.12s. The MIP method achieves convergence in 257 instances with an average time of 13.28s; for the remaining 1272 instances (83.2% of the total instances), the MIP method does not converge after 60s, and at that point the solution quality is still worse than the AD solutions (the objective value is on average 1.02% worse than the AD solutions). Those MIP instances exceeding 60s do not achieve much improvement after running for another 10 min. In terms of solution quality, the AD solutions on average obtain an optimality gap of 3.73% compared to the global optimum of the MIP solutions when MIP converges. These comparisons show that the AD algorithm is an effective and efficient heuristic for solving the bilinear program.

2.5.6 Tests on 118-bus system

Extensive simulation is also conducted on the 118-bus system. The results for this larger system support similar conclusions as shown in the 14-bus system. The 118-bus system has 54 generators of total 7220 MW generation capacity and 273.2 MW/10min system ramping capacity. There are 186 lines with flow limits ranging between 280 MW and 1000 MW, and 91 loads. Total electricity demand is between 2485.7 MW (3:30 am) and 5982.9 MW (8:20 pm) with an average of 4735.0 MW. There are 8 wind farms, each with a capacity $p^{w,max} = 750$ MW. The average total available wind

power at any time is 1882.7 MW, equivalent to 31.5% of the peak demand. All the wind speeds used in simulation are real data collected from [21]. Each robust ED model can still be quickly solved in about 20 seconds in the laptop described before. For the simulation of 35-day rolling horizon with a 10-min interval, we use a computer cluster [39].

Table 5 shows the performance of the deterministic LA-ED and the Rob-ED with dynamic uncertainty sets of lags $L = 6$. From the table, we have the following observations:

(1) Rob-ED reduces the average cost by 43.4% ((15061−8528)/15061) at a properly chosen $\Gamma^w = 1.5$.

(2) Cost std is reduced by 87.7% at $\Gamma^w = 1.5$ and by 93.9% at $\Gamma^w = 2.0$.

(3) The average penalty cost is reduced by 98.4% or 60.7 times at $\Gamma^w = 1.5$ and is almost eliminated at $\Gamma^w = 2.0$. The frequency of penalty is 7.70% by LA-ED, and is reduced to 0.12% and 0.02% by Rob-ED at $\Gamma^w = 1.5$ and 2.0, respectively.

(4) Rob-ED dispatches more thermal and curtails more wind. On average, the thermal generation is up by 12.7% and 18.9%, and the wind generation is down by 24% and 38.9%, at $\Gamma^w = 1.5$ and 2.0, respectively. This can be explained by a similar reasoning given in Section 2.5.1.2, namely that the robust ED dispatches the thermal generation anticipating to a potential large drop of wind in the future, optimally balancing thermal and wind generation in the system.

Comparing to the 14-bus system, the above results for the 118-bus system show a more significant benefit of the proposed Rob-ED model: the average operating cost is cut to almost half of the look-ahead ED, the cost variability is reduced by an order of magnitude, and the shortage events and penalty cost are almost eliminated. Table 5 also shows a Pareto frontier exists for the range of Γ^w between 1.5 to 2.0.

Table 5: Performance of LA-ED and Rob-ED for 118-bus system

	LA-ED	Rob-ED			
Γ^w	0.0	0.5	1.0	1.5	2.0
Cost Avg ($)	15061	12193	8914	8528	9075
Cost Std ($)	38138	30903	14671	4703	2325
Penalty Avg ($)	7775	4835	1214	126	1
Penalty Freq (%)	7.70	4.74	1.45	0.12	0.02
Therm Avg (MW)	2969	3007	3132	3399	3660
Wind Avg (MW)	1758	1723	1602	1336	1075

2.6 Conclusion

In this Chapter, we presented an adaptive multi-period robust ED model and dynamic uncertainty sets for power system economic dispatch under high penetration levels of wind resources. The adaptive multi-period robust ED model mimics the physical dispatch procedure by using a two-stage decision making structure and a rolling-horizon framework. Dynamic uncertainty sets explicitly model the relationship between uncertainties across decision stages and capture the temporal and spatial correlations of wind power output in multiple wind farms: the proposed dynamic uncertainty sets with linear dynamics in this Chapter have general and computationally tractable structure; and the proposed data-driven estimation procedures are easy to implement. We also develop a simulation platform that integrates the optimization engine and data analysis tools for updating uncertainty sets.

Extensive simulation using real wind data shows that the proposed robust ED framework outperforms look-ahead ED models with and without reserves which recently attracted considerable interests in practice, and robust ED models with static uncertainty sets. Both cost efficiency and system reliability are substantially improved. Also, the robust ED model gives an entire Pareto frontier of operating cost and reliability, which provides an informative guideline for choosing uncertainty set parameters and system operating points. The proposed robust ED model and dynamic uncertainty sets are flexible enough to incorporate several extensions, such as

using transformed wind speeds, bids with piecewise linear costs, and including other types of uncertain renewable energy sources.

CHAPTER III

MULTISTAGE ADAPTIVE ROBUST OPTIMIZATION FOR THE UNIT COMMITMENT PROBLEM

3.1 Introduction

Operating large-scale electric power systems is a challenging task that requires adequate decision tools and methodologies for hedging against uncertain factors such as wind and solar power generation, water inflows for hydroplants, electricity demand, transmission line and generator contingencies, and demand response (see e.g. [41, 26, 110]). The unit commitment (UC) problem consists in finding an on/off commitment schedule and generation dispatch levels for generating units in each hour of the next day, in such a way that the total production cost is minimized while electricity demand is met and various physical constraints of generators and the transmission network are satisfied. This is the most critical daily operational problem for large-scale power systems, and it is a difficult optimization problem due to its large scale and discrete nature. It becomes more complicated when wind power and other renewable generation resources are available in large quantities and present significant uncertainty in their availability. How to deal with increasing uncertainty in power systems has been identified by the electricity industry as an urgent challenge (see [47, 57, 26, 110]).

Stochastic programming is an important approach that has been applied to managing uncertainties in the UC problem (e.g., see [80], [108], [101], [88], [89], [98], [27],[102], [82]). These models offer a notable advancement over deterministic methods. However, they also present important computational challenges when dealing

with large-scale power systems. In particular, stochastic programming models require identifying appropriate probability distributions for uncertain parameters such as load and renewable energy generation, which might be difficult; it is also difficult to construct scenario trees that represent high-dimensional stochastic processes; and large scenario trees often lead to computational difficulties. See the work by [46] for an important example on scenario tree generation methods.

Robust optimization is an alternative paradigm for optimization under uncertainty, which has received wide attention and has been applied in several engineering disciplines (e.g., see [5], [9]). Instead of using probability distributions for uncertain parameters, robust optimization models assume that uncertain parameters are realized as elements of a deterministic *uncertainty set*. Given an uncertainty set, the problem consists of finding a solution that is feasible for any realization of the uncertain parameters in this set and also minimizes the worst-case cost. This approach is of particular interest when accurate probability distributions are not available or when uncertain parameters present high dimensionality. Further, the conservativeness of robust solutions can be controlled by the choice of uncertainty sets.

Several robust optimization formulations for the UC problem have been recently proposed. A robust formulation for the contingency constrained UC problem is proposed in [94]. Various adaptive robust UC models dealing with demand and renewable generation uncertainty are studied in [53], [118], [13], and [103]. More specifically, [53] present a robust UC formulation including pumped-storage hydro units under wind uncertainty. [118] present a formulation with demand response under wind uncertainty. [13] present a security constrained robust UC formulation with system reserve requirements under nodal net injection uncertainty, including extensive computational experiments on a real-world power system. [103] present a contingency constrained UC model under uncertain generator and transmission line contingencies. [115] present a hybrid approach that combines stochastic and robust optimization by

weighing expected cost and worst-case cost.

An essential feature of all the above stochastic and robust UC models is that they are *two-stage* models, where the first-stage decision is the on/off commitment decision made in the day-ahead electricity market, while the second-stage decision is the real-time dispatch decision for the entire scheduling horizon. The work in [116] presents a three-stage robust UC model, which has UC decisions in the first stage and dispatch decisions in the second stage, and then has uncertain demand response after dispatch decisions. This decision-making structure is converted to a two-stage robust model. The crucial assumption of two-stage models is that the second-stage decision is made with the full knowledge of realized uncertain parameters over the entire scheduling horizon. However, in reality, power systems are operated sequentially, where generation dispatch at each hour can only depend on the information of realized uncertain parameters up to that hour. In other words, dispatch decisions are *non-anticipative*. Two-stage stochastic and robust UC models ignore this.

In this Chapter, we demonstrate the importance of considering non-anticipativity constraints in power system operations and present a *multistage* adaptive robust optimization model for the UC problem, where the commitment decisions are selected here-and-now as done in the day-ahead electricity market, and the dispatch decision for each hour of the next day is the wait-and-see decision respecting non-anticipativity for the sequential revelation of uncertainty. We also address the computational challenge presented by the multistage robust UC model. To make it computationally tractable, we consider approximation schemes with decision rules, in particular, we use *affine policies* for the dispatch decisions, where generators' dispatch levels are affine functions of uncertain load.

The affinely adjustable robust optimization approach has attracted considerable attention since the seminal paper of [7]. Most of the existing works focus on studying multistage convex optimization problems with relatively simple and well-structured

constraints, such as multi-period inventory problems studied in [12, 40, 44] or multistage stochastic linear programs in [61]. These models can be transformed to deterministic counterparts through duality theory and solved by existing algorithms for convex programs, see e.g. [6, 61]. Another direction of research is to extend affine policies to more general decision rules, such as in [23], [24], and [38].

Previous applications of affine policies for power system operations were carried out by [104, 105], who considered a stochastic optimization model for the economic dispatch problem with storage, where the UC binary decision is assumed to be fixed. These works have been recently extended to incorporate binary UC decisions [106]. Another application of affine policies for power system operations was developed by [51] who considered the dispatch of automatic generation control units under uncertain renewable energy outputs, with fixed UC decisions. Our work is done independently from these works. And the crucial differences of our approach with respect to these references include the proposal of a multistage robust UC model with simplified affine policies, the analysis of the relationship between the multistage and two-stage robust UC models, and the development of an algorithm based on constraint generation that allows the efficient solution of large-scale instances of the problem under a high-dimensional uncertainty set. An interesting analysis comparing two-stage and multistage robust formulations is presented by [72], for an economic dispatch problem with one bus and one generator. However, no details on the multistage model are provided. The author also introduces the idea of state-space representable uncertainty sets, which can be used for modeling temporal dependencies in the uncertain parameters.

The proposed multistage robust UC problem in this Chapter presents several challenges that make existing methodologies not directly applicable. In particular, the multistage robust UC model is a large-scale mixed-integer optimization problem involving a large number of complicated constraints. Due to the mixed-integer

decisions, convex optimization based modeling and solution methods cannot be applied. Furthermore, due to its very large scale, applying even the basic affine policies in a straightforward way is not computationally viable and the duality-based approach leads to reformulations with exceedingly large dimensions. To deal with these challenges, we propose new solution ideas. More specifically, instead of using more general decision rules, we descend the complexity ladder and use *simplified affine policies* through properly aggregating uncertain parameters in the dependency structure of the affine policy. The resulting multistage UC formulation has a reduced dimensionality and a structure that we can exploit. We design a solution method based on constraint generation and employ several algorithmic improvements that make the multistage robust UC problem efficiently solvable even for large-scale instances.

We conduct a thorough computational study with extensive numerical experiments on the performance of the proposed algorithm, the quality of simplified affine policies, their worst-case and average-case performance, and comparison with existing deterministic and two-stage robust UC models. The computational results show that the proposed algorithm can effectively solve the multistage robust UC model within a time frame reasonable for the day-ahead operation of large-scale power systems. The performance of the proposed multistage robust UC model demonstrates its ability to significantly reduce operational costs and at the same time improve system reliability, as we show in computational experiments where we compare this approach with the existing deterministic and two-stage robust UC models.

The contributions of this Chapter can be summarized as follows.

1. This Chapter presents a multistage adaptive robust optimization model for the UC problem under significant uncertainty in nodal net loads, respecting non-anticipativity in the dispatch process. We have also proposed a new robust dispatch model utilizing the affine policy obtained from the multistage robust UC model.

2. This Chapter discusses the solution concept of simplified affine policies in multistage robust optimization and demonstrates its effectiveness in power system operations.

3. This Chapter proposes an efficient solution algorithm based on constraint generation with various algorithmic improvements for solving large-scale multistage robust UC models with affine policy. Several aspects of the algorithm are also applicable to general large-scale robust optimization problems with mixed-integer variables.

4. This Chapter provides an extensive computational study of the proposed multistage robust UC model on medium and large-scale power systems. Comparison with existing deterministic and two-stage robust UC models demonstrates the potential of the proposed approach in reducing operational cost, increasing system reliability, and managing system flexibility.

The remainder of this Chapter proceeds as follows. Section 3.2 presents the deterministic and two-stage robust UC models and discusses their limitations. Section 3.3 proposes the multistage robust UC model and introduces the concept of simplified affine policies. Section 3.4 presents a traditional method based on duality and a constraint generation framework for solving robust optimization problems. Section 3.5 proposes several algorithmic improvements. Section 3.6 presents a multifaceted computational study of the performance of the proposed approach. Section 3.7 concludes this Chapter. All the proofs, unless given in the main body of this Chapter, are collected in Appendix 3.8.

3.2 Non-Causal UC models and Their Limitations

In this section, we discuss the deterministic UC and the recently developed two-stage robust UC models. We shall call the two-stage robust UC model a *non-causal* UC

model, because the decisions in this model depend on future information of uncertainty and thus do not respect non-anticipativity in the physical process of dispatching generators. We show important issues with non-causal UC formulations. It serves as the motivation for the development of multistage robust UC models.

3.2.1 Deterministic Unit Commitment

Consider the deterministic UC model below.

$$\min_{\boldsymbol{x},\boldsymbol{u},\boldsymbol{v},\boldsymbol{p}} \quad \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}_g}\left(G_i x_i^t + S_i u_i^t\right) + \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}_g} C_i p_i^t \tag{16a}$$

$$\text{s.t.} \quad x_i^t,\, u_i^t,\, v_i^t \in \{0,1\} \qquad \forall\, i\in\mathcal{N}_g,\, t\in\mathcal{T} \tag{16b}$$

$$x_i^t - x_i^{t-1} = u_i^t - v_i^t \qquad \forall\, i\in\mathcal{N}_g,\, t\in\mathcal{T} \tag{16c}$$

$$\sum_{\tau=t}^{t+UT_i-1} x_i^\tau \geq UT_i u_i^t \qquad \forall\, i\in\mathcal{N}_g,\, t\in\{1,2,\ldots,\,T-UT_i+1\} \tag{16d}$$

$$\sum_{\tau=t}^{t+DT_i-1} (1-x_i^\tau) \geq DT_i v_i^t \qquad \forall\, i\in\mathcal{N}_g,\, t\in\{1,2,\ldots,T-DT_i+1\} \tag{16e}$$

$$\sum_{\tau=t}^{T}(x_i^\tau - u_i^t) \geq 0 \qquad \forall\, i\in\mathcal{N}_g,\, t\in\{T-UT_i+1,\ldots,T\} \tag{16f}$$

$$\sum_{\tau=t}^{T}(1-x_i^\tau - v_i^t) \geq 0 \qquad \forall\, i\in\mathcal{N}_g,\, t\in\{T-DT_i+1,\ldots,T\} \tag{16g}$$

$$p_i^{min} x_i^t \leq p_i^t \leq p_i^{max} x_i^t \qquad \forall\, i\in\mathcal{N}_g,\, t\in\mathcal{T} \tag{16h}$$

$$-RD_i x_i^t - SD_i v_i^t \leq p_i^t - p_i^{t-1} \leq RU_i x_i^{t-1} + SU_i u_i^t \quad \forall\, i\in\mathcal{N}_g,\, t\in\mathcal{T} \tag{16i}$$

$$-f_l^{max} \leq \boldsymbol{\alpha}_l^T(\boldsymbol{B}^p\boldsymbol{p}^t - \boldsymbol{B}^d\boldsymbol{d}^t) \leq f_l^{max} \quad \forall\, t\in\mathcal{T},\, l\in\mathcal{N}_l \tag{16j}$$

$$\sum_{i\in\mathcal{N}_g} p_i^t = \sum_{j\in\mathcal{N}_d} d_j^t \qquad \forall\, t\in\mathcal{T}, \tag{16k}$$

where $\mathcal{N}_g, \mathcal{N}_d, \mathcal{N}_l, \mathcal{T}$ denote the sets of generators, nodes with net load, transmission lines, and time periods, respectively, and N_g, N_d, N_l, T are their cardinalities; x_i^t, u_i^t, v_i^t and p_i^t are respectively the on/off, start-up, shut-down, and the generation dispatch level decisions of generator i at time t; G_i, S_i, C_i are the no-load cost, start-up cost, and variable cost of generator i; DT_i and UT_i are the minimum-down and minimum-up times of generator i; p_i^{min} and p_i^{max} are the minimum and maximum generation levels of generator i; RD_i and RU_i are the ramp-down and ramp-up rates of generator i, and SD_i and SU_i are the ramp rates when generator i shuts down and turns on; $\boldsymbol{B}^p$ and $\boldsymbol{B}^d$ are the incidence matrices for generators and loads; $\boldsymbol{\alpha}_l$ and f_l^{max} are the generation shift factor and the flow limit for line l, respectively; d_j^t is the net load at node j and time t. In this Chapter, nodal net load is defined as the nodal demand minus the total renewable generation such as wind and solar power connected to the same node, which is an uncertain quantity due to the uncertainty of wind and solar power generation. The objective (16a) consists of minimizing the sum of commitment costs (including no-load and start-up costs) and dispatch costs (assumed to be linear here but can be replaced with a piecewise linear function without changing the linearity of the problem). Eq. (16c) corresponds to start-up and shut-down constraints. Eq. (16d)-(16g) corresponds to minimum up and down time constraints. Constraints (16h) enforce minimum and maximum generation capacity limits when generators are on, and no generation when they are off. Constraints (16i) enforce ramping up and down limits. Constraints (16j) enforce transmission line limits. Constraints (16k) enforce energy balance at a system level. The model can also be extended to include reserve decisions and related constraints, which are omitted here for simplicity. The formulation in (16d)-(16g) follows [79].

3.2.2 Two-stage Adaptive Robust Unit Commitment

To deal with uncertainties in the nodal net electricity loads, the following two-stage adaptive robust UC model has been proposed (e.g. see [53], [118], [13]):

$$\min_{\boldsymbol{x}\in X} \left\{ F(\boldsymbol{x}) + \max_{\boldsymbol{d}\in\mathcal{D}} \min_{\boldsymbol{p}\in\Omega(\boldsymbol{x},\boldsymbol{d})} c(\boldsymbol{p}) \right\}, \tag{17}$$

where $\boldsymbol{x}$ denotes all the commitment related binary variables (x_i^t, u_i^t, v_i^t in the deterministic UC model (16)), $\boldsymbol{d}$ is the vector of net load at all nodes and all time periods, $\boldsymbol{p}$ is the vector of dispatch variables, set X is the feasible region of the commitment decisions defined by Eq. (16b)-(16g), $\mathcal{D}$ is the uncertainty set of net loads, $\Omega(\boldsymbol{x}, \boldsymbol{d})$ is the feasible region of the dispatch variables parameterized by the commitment decisions and realized net load, as defined in Eq. (16h)-(16k), and

$$F(\boldsymbol{x}) = \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}_g} \left(G_i x_i^t + S_i u_i^t \right), \quad c(\boldsymbol{p}) = \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}_g} C_i p_i^t.$$

In this Chapter, we use the following budget uncertainty set:

$$\mathcal{D}^t = \left\{ \boldsymbol{d}^t = (d_1^t, \ldots, d_{N_d}^t) : \sum_{j\in\mathcal{N}_d} \frac{|d_j^t - \overline{d}_j^t|}{\hat{d}_j^t} \le \Gamma\sqrt{N_d}, \;\; d_j^t \in [\overline{d}_j^t - \Gamma\hat{d}_j^t,\, \overline{d}_j^t + \Gamma\hat{d}_j^t] \;\; \forall\, j \in \mathcal{N}_d \right\}. \tag{18}$$

Notice that d_j^t lies in an interval centered around the nominal value $\overline{d}_j^t$ within a deviation denoted by $\Gamma\hat{d}_j^t$. The budget constraint with budget $\Gamma\sqrt{N_d}$ controls the size of the uncertainty set, where Γ represents the conservativeness of the model. For $\Gamma = 0$ we have $\mathcal{D}^t = \{\overline{\boldsymbol{d}}^t\}$, i.e., the uncertainty set only contains the nominal net load vector and the two-stage robust UC model (17) becomes the deterministic UC model (16). As Γ increases, more net load vectors are contained in the uncertainty set. The square root $\sqrt{N_d}$ scaling is motivated by a central limit theorem type argument, where the standard deviation of the aggregated randomness scales proportionally to the square root of the number of random variables (see [13]). Robust constraints using uncertainty set (18) also guarantees a probabilistic feasibility condition (see

[22]). We define $\mathcal{D} = \prod_{t \in \mathcal{T}} \mathcal{D}^t$ as the uncertainty set for the net load trajectory $\boldsymbol{d}$ over the entire scheduling horizon. With this choice, notice the separability of $\mathcal{D}$ over time periods, i.e., the temporal independency.

As seen from the above two-stage robust UC model, the dispatch decision $\boldsymbol{p}$ in the inner minimization problem over $\Omega(\boldsymbol{x}, \boldsymbol{d})$ is made with perfect knowledge of the realization of uncertain net loads $\boldsymbol{d}$ over the entire scheduling horizon. In reality, system operators only have perfect information about uncertain parameters that are realized up to the current operating time. The key questions are: What is the consequence of assuming the full knowledge of nodal net loads in the dispatch process? How to properly tackle the sequential nature of this process?

3.2.3 Example that Illustrates the Limitations of Non-causal UC Models

We present a simple example to illustrate that the UC solution from the two-stage robust UC model can lead to infeasibility in the real-time dispatch.

Example 1. The system has 2 buses, A and B, and two periods, so $T = 2$. Each bus has a conventional generator. The transmission line has a flow limit of 1 unit of power. The ramp rates of both generators are also 1 unit of power, i.e., $RU_A = RD_A = R_A = 1$, and $RU_B = RD_B = R_B = 1$. The initial generation levels of the two generators are at 12, i.e., $p_A^0 = p_B^0 = 12$ at $t = 0$.

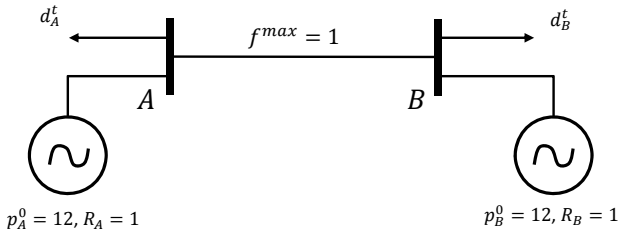

Figure 10: Simple two-bus system to illustrate the limitation of non-causal UC models.

The uncertainty sets for nodal net loads (d_A^t, d_B^t) at $t = 1, 2$ are given as follows:

$$\mathcal{D}^1 = \left\{(d_A^1, d_B^1) = (12, 12)\right\}, \text{and } \mathcal{D}^2 = \left\{(d_A^2, d_B^2) :\ d_A^2 \in [10, 15],\ \ d_B^2 \in [10, 15],\ d_A^2 + d_B^2 = 25\right\}.$$

That is, the first period loads are deterministic with power level of 12 at each bus, and the net loads in the second period are uncertain, but the total net load is known to be 25. Denote $\mathcal{D} = \mathcal{D}_1 \times \mathcal{D}_2$.

Claim 1. *The two-stage robust UC model* (17) *is feasible for the system in Example 1.*

Proof. Proof: Consider $\boldsymbol{x}_{2S} = ((x_A^1, x_B^1), (x_A^2, x_B^2)) = ((1, 1), (1, 1))$. To prove that $\boldsymbol{x}_{2S}$ is feasible for the two-stage robust UC model, we construct a feasible dispatch policy. In particular, for any $\boldsymbol{d} = ((d_A^1, d_B^1), (d_A^2, d_B^2)) \in \mathcal{D}$, consider the following policy:

$$p_A^1(\boldsymbol{d}) = 12 + (2/5)(d_A^2 - 12.5),\ \ p_B^1(\boldsymbol{d}) = 12 - (2/5)(d_A^2 - 12.5), \tag{19a}$$

$$p_A^2(\boldsymbol{d}) = 12.5 + (3/5)(d_A^2 - 12.5),\ \ p_B^2(\boldsymbol{d}) = 12.5 - (3/5)(d_A^2 - 12.5). \tag{19b}$$

From (19), we can see that for $t = 1$, $p_A^1(\boldsymbol{d}) + p_B^1(\boldsymbol{d}) = 24$ for all $\boldsymbol{d} \in \mathcal{D}$, so energy balance is respected. By the definition of the uncertainty sets, we have $p_A^1(\boldsymbol{d}), p_B^1(\boldsymbol{d}) \in [11, 13]$ for all $\boldsymbol{d} \in \mathcal{D}$, so the ramping constraints from the initial states ($\boldsymbol{p}^0 = (12, 12)$) are respected. Furthermore, $p_A^1(\boldsymbol{d}) - d_A^1 = p_A^1(\boldsymbol{d}) - 12 \in [-1, 1]$ for all $\boldsymbol{d} \in \mathcal{D}$, so transmission constraints are respected. Similarly for $t = 2$, we have $p_A^2(\boldsymbol{d}) + p_B^2(\boldsymbol{d}) = 25$ for all $\boldsymbol{d} \in \mathcal{D}$, so energy balance is satisfied. Since $d_A^2 \in [10, 15]$, we can see that $p_A^2(\boldsymbol{d}) - p_A^1(\boldsymbol{d}) = 0.5 + (1/5)(d_A^2 - 12.5) \in [0, 1]$ and $p_B^2(\boldsymbol{d}) - p_B^1(\boldsymbol{d}) = 0.5 - (1/5)(d_A^2 - 12.5) \in [0, 1]$, hence ramping constraints are respected. Finally, $p_A^2(\boldsymbol{d}) - d_A^2 = 5 - (2/5)d_A^2 \in [-1, 1]$, so transmission constraints are respected. Therefore, $\boldsymbol{p}(\boldsymbol{d})$ given in (19) satisfies all constraints in (17), thus $\boldsymbol{x}_{2S}$ is feasible for the two-stage robust UC model. □

Notice that the dispatch policy (19) is non-causal, because the dispatch decision at $t = 1$ depends on the uncertainty realization at $t = 2$. If this UC solution is implemented, the real-time dispatch under this UC solution can be infeasible, as shown in the following result.

Claim 2. *Let $\boldsymbol{x}^*_{2S}$ be the optimal UC solution of the two-stage robust UC model for Example 1. Under $\boldsymbol{x}^*_{2S}$, there does not exist any feasible dispatch policy that respects time causality, i.e. where $\boldsymbol{p}^1(\cdot)$ does not depend on $\boldsymbol{d}^2$.*

Proof. Proof Notice that $\boldsymbol{x}^*_{2S} = ((x^1_A, x^1_B), (x^2_A, x^2_B)) = ((1,1),(1,1))$ is the optimal solution of the two-stage robust UC for the system in Example 1, since keeping both generators online is the only candidate solution to satisfy net load.

Now consider the real-time sequential operation under this commitment decision $\boldsymbol{x}^*_{2S}$, where the causal dispatch policy at t can only depend on the information available up to t. We want to show that there does not exist any causal dispatch policy that can make the system feasible for all net load vectors in the uncertainty set. For this, we need to show that there is no $\boldsymbol{p}^1(\boldsymbol{d}^1)$ such that, for all $\boldsymbol{d}^2 \in \mathcal{D}^2$, there always exists a feasible $\boldsymbol{p}^2(\boldsymbol{d}^1, \boldsymbol{d}^2)$ at $t = 2$.

Since $\boldsymbol{d}^1$ is fixed at $(12, 12)$, we write $\boldsymbol{p}^1(\boldsymbol{d}^1)$ as $\boldsymbol{p}^1 = (p^1_A, p^1_B)$ for brevity. Notice that due to the energy balance constraint, we must have $p^1_A + p^1_B = 24$, and due to the ramping capacity and transmission constraints, we must have $p^1_A, p^1_B \in [11, 13]$. Suppose we choose $p^1_A \leq 12$. Then take $\boldsymbol{d}^2 = (15, 10)$ from the uncertainty set $\mathcal{D}^2$. Due to ramping constraints, we must have $p^2_A \leq 13$. However, it is impossible to satisfy energy balance at location A, because the transmission limit is 1. Similarly, if we choose $p^1_A \geq 12$, the adversary can take $\boldsymbol{d}^2 = (10, 15) \in \mathcal{D}^2$, which leads to the impossibility of satisfying net load at location B. This means that no matter what $\boldsymbol{p}^1$ we choose to satisfy the constraints at $t = 1$, there always exists a $\boldsymbol{d}^2 \in \mathcal{D}^2$ so that the constraints at $t = 2$ cannot be satisfied. □

With this result, we see that the two-stage robust UC decision $\boldsymbol{x}^*_{2S}$ can lead to infeasibility in the real-time dispatch problem. This simple example demonstrates that when the transmission and generation ramping capability is limited, the two-stage robust UC model can make an infeasible problem appear to be feasible. When such a UC solution is implemented, the real-time operation can become infeasible under uncertain parameters realized within the uncertainty set. Also notice that, if we add expensive generators at each bus in Example 1, we can obtain a system where the multistage robust UC model produces a UC solution under which feasible real-time dispatch is guaranteed, while the optimal commitment solution from the two-stage robust UC solution again leads to infeasibility in real-time operation.

With high penetration of renewable energy resources, power systems frequently experience fast swings in net loads, which pushes the generators toward the regime of limited ramping capability, which motivates us to consider multistage robust UC models.

3.3 Multistage Adaptive Robust UC and Simplified Affine Policy

In this section, we first propose the multistage robust UC model and give a theoretical analysis on the relationship between the two-stage and multistage robust UC models. Then, we introduce affine dispatch policies and the concept of simplified affine policies.

3.3.1 Multistage Adaptive Robust UC Model

In the operation of power systems, the commitment decision $\boldsymbol{x}$ is made several hours before the observations of uncertain net loads, and then the dispatch decisions are sequentially optimized in real time with observations of realized uncertainty up to the operating hour. To faithfully model this process, the dispatch decision $\boldsymbol{p}^t$ at time t in the UC model should depend on the history of net load $\boldsymbol{d}^{[t]} \triangleq (\boldsymbol{d}^1, ..., \boldsymbol{d}^t)$. Based on

this requirement, we formulate the following multistage adaptive robust UC model.

$$\min_{\boldsymbol{x},\boldsymbol{u},\boldsymbol{v},\boldsymbol{p}(\cdot)} \left\{ \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}_g} \left(G_i x_i^t + S_i u_i^t\right) + \max_{\boldsymbol{d}\in\mathcal{D}} \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}_g} C_i\, p_i^t(\boldsymbol{d}^{[t]}) \right\} \tag{20a}$$

s.t. Constraints (16b)-(16g) for $(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v})$

$$p_i^{min} x_i^t \le p_i^t(\boldsymbol{d}^{[t]}) \le p_i^{max} x_i^t \qquad \forall\, \boldsymbol{d}\in\mathcal{D},\, i\in\mathcal{N}_g,\, t\in\mathcal{T} \tag{20b}$$

$$-RD_i x_i^t - SD_i v_i^t \le p_i^t(\boldsymbol{d}^{[t]}) - p_i^{t-1}(\boldsymbol{d}^{[t-1]}) \le RU_i x_i^{t-1} + SU_i u_i^t \qquad \forall \boldsymbol{d}\in\mathcal{D},\, i\in\mathcal{N}_g,\, t\in\mathcal{T} \tag{20c}$$

$$-f_l^{max} \le \boldsymbol{\alpha}_l^T(\boldsymbol{B}^p \boldsymbol{p}^t(\boldsymbol{d}^{[t]}) - \boldsymbol{B}^d \boldsymbol{d}^t) \le f_l^{max} \quad \forall\, \boldsymbol{d}\in\mathcal{D},\, t\in\mathcal{T},\, l\in\mathcal{N}_l \tag{20d}$$

$$\sum_{i\in\mathcal{N}_g} p_i^t(\boldsymbol{d}^{[t]}) = \sum_{j\in\mathcal{N}_d} d_j^t \qquad \forall\, \boldsymbol{d}\in\mathcal{D},\, t\in\mathcal{T}. \tag{20e}$$

The crucial feature of this formulation is the expression $p_i^t(\boldsymbol{d}^{[t]})$, which makes the generation of unit i at time t a function of net load uncertainty realized up to time t, thus respecting non-anticipativity. Constraints (20b)-(20e) enforce generation limits, ramping capacities, transmission line capacities, and energy balance, for any realization of $\boldsymbol{d}\in\mathcal{D}$. Note that binary decisions $\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}$ are not adaptive, they are decided "here-and-now" before observing any uncertainty.

The multistage decision making structure of (20) can be equivalently represented in the following nested formulation, using the separability of the uncertainty set over time periods:

$$\min_{(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v})\in X} \left\{ \boldsymbol{G}^\top \boldsymbol{x} + \boldsymbol{S}^\top \boldsymbol{u} + \max_{\boldsymbol{d}^1\in\mathcal{D}^1} \min_{\boldsymbol{p}^1\in\Omega_1(\boldsymbol{x},\boldsymbol{d}^1,\boldsymbol{p}^0)} \left\{ \boldsymbol{C}^\top \boldsymbol{p}^1 + \cdots + \max_{\boldsymbol{d}^T\in\mathcal{D}^T} \min_{\boldsymbol{p}^T\in\Omega_T(\boldsymbol{x},\boldsymbol{d}^T,\boldsymbol{p}^{T-1})} \boldsymbol{C}^\top \boldsymbol{p}^T \right\} \right\}, \tag{21}$$

where $\Omega_t(\boldsymbol{x}, \boldsymbol{d}^t, \boldsymbol{p}^{t-1}) \triangleq \left\{ \boldsymbol{p}^t :\ \text{(16h)-(16k) are satisfied} \quad \forall\, i\in\mathcal{N}_g \right\}$. Notice that the feasible region $\Omega_t(\boldsymbol{x}, \boldsymbol{d}^t, \boldsymbol{p}^{t-1})$ of the dispatch decision at stage t depends on previous stage $t-1$'s dispatch levels $\boldsymbol{p}^{t-1}$ and stage t's realized demand $\boldsymbol{d}^t$. Due to discrete decision variables and the large scale of the formulation, numerical solution of the multistage robust UC model ((20) or (21)) presents a major computational challenge.

In the following, we first make further discussion on the relation between the two-stage and multistage models, then propose approximate decision rules and tractable solution methods for solving the multistage robust UC model.

3.3.2 Discussion on Two-Stage and Multistage Robust UC Models

The key difference between the two-stage and multistage models is that the multistage robust UC provides a *causal* policy $\boldsymbol{p}^t(\boldsymbol{d}^{[t]})$, which only relies on information observed up to the respective time period when the dispatch decision is made. Clearly, the two-stage robust UC model lacks this property. It turns out that, when the system is not constrained by its ramping capability, i.e., the generators have enough ramping capacity to follow the rapidly varying wind, the two-stage robust UC model is equivalent to the multistage model.

Proposition 3. *Consider the two-stage robust UC* (17) *and the multistage robust UC* (20)*, where the uncertainty set is given by Eq.* (18)*. If ramping constraints* (20c) *are neglected, the two-stage robust UC* (17) *and the multistage robust UC* (20) *have the same optimal objective value and a same optimal UC solution.*

The proof follows from the fact that the dispatch component of both problems can be separated into T non-coupled problems when there are no ramping constraints. Please see Appendix 3.8 for details.

Proposition 3 suggests that the multistage robust UC model is important precisely when the system's ramping capability is a limited resource, which is the case for power systems with a high penetration of uncertain wind and solar power generation.

3.3.3 Affine Multistage Robust UC

To computationally solve the proposed multistage robust UC model (20), we propose to consider approximation schemes using linear decision rules. In particular, to make the problem tractable, we restrict the dispatch decision $\boldsymbol{p}^t(\cdot)$ to have the form of an

affine function as

$$p_i^t(\boldsymbol{d}^{[t]}) = w_i^t + \sum_{j\in\mathcal{N}_d}\sum_{s\in[1:t]} W_{itjs} d_j^s, \tag{22}$$

where $[1:t] \triangleq \{1, ..., t\}$ and (w_i^t, W_{itjs}) are the coefficients of the affine policy. It is important to notice that the affine policy (22) automatically respects non-anticipativity. Using this affine dispatch policy, the multistage robust UC model has the following form

$$\min_{\boldsymbol{x},\boldsymbol{u},\boldsymbol{v},z,\boldsymbol{w},\boldsymbol{W}} \quad \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}_g}\left(G_i x_i^t + S_i u_i^t\right) \; + \; z \tag{23a}$$

s.t. Constraints (16b)-(16g) for $(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v})$

$$\sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}_g} C_i\left(w_i^t + \sum_{j\in\mathcal{N}_d}\sum_{s\in[1:t]} W_{itjs} d_j^s\right) \le z \qquad \forall\, \boldsymbol{d}\in\mathcal{D} \tag{23b}$$

$$p_i^{min} x_i^t \le w_i^t + \sum_{j\in\mathcal{N}_d}\sum_{s\in[1:t]} W_{itjs} d_j^s \le p_i^{max} x_i^t \qquad \forall\, \boldsymbol{d}\in\mathcal{D},\, i\in\mathcal{N}_g,\, t\in\mathcal{T} \tag{23c}$$

$$\left(w_i^t + \sum_{j\in\mathcal{N}_d}\sum_{s\in[1:t]} W_{itjs} d_j^s\right) - \left(w_i^{t-1} + \sum_{j\in\mathcal{N}_d}\sum_{s\in[1:t-1]} W_{i,t-1,js} d_j^s\right) \ge -RD_i x_i^t - SD_i v_i^t \qquad \forall\, \boldsymbol{d}\in\mathcal{D},\, i\in\mathcal{N}_g,\, t\in\mathcal{T} \tag{23d}$$

$$\left(w_i^t + \sum_{j\in\mathcal{N}_d}\sum_{s\in[1:t]} W_{itjs} d_j^s\right) - \left(w_i^{t-1} + \sum_{j\in\mathcal{N}_d}\sum_{s\in[1:t-1]} W_{i,t-1,js} d_j^s\right) \le RU_i x_i^{t-1} + SU_i u_i^t \qquad \forall\, \boldsymbol{d}\in\mathcal{D},\, i\in\mathcal{N}_g,\, t\in\mathcal{T} \tag{23e}$$

$$-f_l^{max} \le \sum_m\sum_{i\in\mathcal{N}_g} \alpha_{lm} B_{mi}^p\left(w_i^t + \sum_{j\in\mathcal{N}_d}\sum_{s\in[1:t]} W_{itjs} d_j^s\right) - \sum_m\sum_{j\in\mathcal{N}_d} \alpha_{lm} B_{mj}^d d_j^t \;\le\; f_l^{max} \qquad \forall\, \boldsymbol{d}\in\mathcal{D},\, t\in\mathcal{T},\, l\in\mathcal{N}_l \tag{23f}$$

$$\sum_{i\in\mathcal{N}_g}\left(w_i^t + \sum_{j\in\mathcal{N}_d}\sum_{s\in[1:t]} W_{itjs} d_j^s\right) = \sum_{j\in\mathcal{N}_d} d_j^t \qquad \forall \boldsymbol{d}\in\mathcal{D}, t\in\mathcal{T}. \tag{23g}$$

We have created variable z to denote the worst-case dispatch cost in constraint (23b). Constraints (23c)-(23g) correspond to (20b)-(20e), obtained by replacing $p_i^t(\boldsymbol{d}^{[t]})$ with the affine policy (22). Note that constraints (23b)-(23g) are robust constraints that should hold for all $\boldsymbol{d} \in \mathcal{D}$. We call (23) the affine multistage robust UC model.

3.3.4 Simplified Affine Policies

In the affine policy (22), the dispatch decision $p_i^t(\boldsymbol{d}^{[t]})$ of generator i at time t depends on the entire history of realized net load at every node and every time period up to t. This full affine dependency requires defining a large number of W_{itjs} variables, which can quickly lead to scalability issues in large-scale power systems.

To make the affine multistage robust UC model (23) a practical decision tool for the operation of large-scale power systems, we introduce further restrictions on the affine policy form. In particular, we consider affine policies with *simplified* structures by limiting the degrees of freedom in W_{itjs}. There are several ways to do this: we can restrict $\boldsymbol{p}^t(\cdot)$ to only depend on the most recently revealed information at time t, rather than on the whole history; we can partition time periods into peak-load, medium-load, and low-load periods and assume affine policies in each period have the same form; we can also partition the transmission network into zones and make generators' dispatch policies depend on the aggregated load in each zone. We use the following two very simple policies:

$$p_i^t(\boldsymbol{d}^{[t]}) = w_i^t + W_i \sum_{j \in \mathcal{N}_d} d_j^t \qquad \forall i \in \mathcal{N}_g, t \in \mathcal{T} \tag{24}$$

$$p_i^t(\boldsymbol{d}^{[t]}) = w_i^t + W_{it} \sum_{j \in \mathcal{N}_d} d_j^t \qquad \forall i \in \mathcal{N}_g, t \in \mathcal{T}. \tag{25}$$

We call (24) the W_i-policy, where the coefficients W_i of the affine policy only depend on generators but not on time, and the dispatch level of each generator at time t depends on the total load in the system at time t. Eq. (25) presents a finer policy, which we call the W_{it}-policy, where the coefficients W_{it} of the affine policy can change over

time. Surprisingly, it will be shown that these two very simplified affine policies are already quite powerful and produce close-to-optimal performance for the multistage robust UC model. We also want to remark that a static policy, i.e. $p_i^t(\boldsymbol{d}^{[t]}) = w_i^t$, is the simplest (and trivial) form of an affine policy, however, with this choice it becomes impossible to satisfy energy balance equality over all net load vectors in the uncertainty set. This shows that the simpler but not the simplest affine policies work and the non-trivial affine dependence in the dispatch policy is very important.

3.4 Basic Algorithmic Framework

In this section, we discuss basic solution methods for the affine multistage robust UC problem (23). We first discuss the traditional approach using duality theory and point out its limitation in solving large-scale robust optimization problems. Then, we present a constraint generation framework as the basis for further algorithmic improvements developed in this Chapter. Then, we close this section with some discussion.

3.4.1 Duality-Based Approach

The robust constraints in (23b)-(23f) have the following structure:

$$\boldsymbol{c}(\boldsymbol{W})^\top \boldsymbol{d} \leq h(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}, z) \quad \forall \boldsymbol{d} \in \mathcal{D}, \tag{26}$$

where $\boldsymbol{c}(\boldsymbol{W})$ and $h(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}, z)$ are affine functions of the respective decision variables. To simplify notations, we write (26) as $\boldsymbol{c}^\top \boldsymbol{d} \leq h$ for all $\boldsymbol{d} \in \mathcal{D}$. This robust constraint can be reformulated by using linear programming duality. In particular, (26) is equivalent to $\max_{\boldsymbol{d} \in \mathcal{D}} \boldsymbol{c}^\top \boldsymbol{d} \leq h$. Since the uncertainty set $\mathcal{D}$ is a polytope, the maximization problem always attains a finite optimum, therefore, the maximization problem can be replaced by the dual minimization problem. Thus, (26) is equivalent to $\min_{\boldsymbol{\pi} \in \Pi(\boldsymbol{c})} \boldsymbol{b}^\top \boldsymbol{\pi} \leq h$, where $\boldsymbol{b}$ comes from the definition of $\mathcal{D}$ and $\Pi(\boldsymbol{c})$ is a polyhedron that depends on $\boldsymbol{c}$. With this, (26) is further equivalent to the existence of

$\boldsymbol{\pi} \in \Pi(\boldsymbol{c})$ such that $\boldsymbol{b}^\top \boldsymbol{\pi} \leq h$. In this way, (26) is reformulated as a finite number of linear constraints involving dual variables. This duality-based approach is general and widely used in reformulating robust constraints, see the book by [5]. For our problem, the deterministic counterpart of (26) with uncertainty set (18) is given below.

Proposition 4. *The robust constraint* $\boldsymbol{c}^\top \boldsymbol{d} = \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{N}_d} c_j^t d_j^t \leq h \;\; \forall\, \boldsymbol{d} \in \mathcal{D}$*, where* $\mathcal{D}$ *is given by* (18)*, is equivalent to the existence of a vector of dual variables* $\boldsymbol{\pi}$ *that satisfies the following linear constraints:*

$$\sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{N}_d} \left[\overline{d}_j^t \pi_{jt}^1 - \overline{d}_j^t \pi_{jt}^2 + (\Gamma \hat{d}_j^t - \overline{d}_j^t)\pi_{jt}^3 + (\Gamma \hat{d}_j^t + \overline{d}_j^t)\pi_{jt}^4\right] + \sum_{t\in\mathcal{T}} \Gamma\sqrt{N_d}\pi_t^5 \leq h \quad \text{(27a)}$$

$$\pi_{jt}^1 - \pi_{jt}^2 - \pi_{jt}^3 + \pi_{jt}^4 = c_j^t \quad \forall\, j \in \mathcal{N}_d,\, t \in \mathcal{T} \quad \text{(27b)}$$

$$-\hat{d}_j^t \pi_{jt}^1 - \hat{d}_j^t \pi_{jt}^2 + \pi_t^5 = 0 \quad \forall\, j \in \mathcal{N}_d,\, t \in \mathcal{T} \quad \text{(27c)}$$

$$\pi_{jt}^1, \pi_{jt}^2, \pi_{jt}^3, \pi_{jt}^4, \pi_t^5 \geq 0 \qquad \forall\, j \in \mathcal{N}_d,\, t \in \mathcal{T}. \quad \text{(27d)}$$

Each robust constraint in (23b)-(23f) can be replaced by a set of equivalent deterministic constraints defined in (27a)-(27d) for the corresponding $\boldsymbol{c}$ and h. Notice that we need to introduce a respective vector $\boldsymbol{\pi}$ of dual variables for each of these robust constraints. The size of the resulting MIP reformulation is very large. In the affine multistage robust UC (23), there are $1 + 2T(2N_g + N_l)$ robust constraints, each requiring a vector $\boldsymbol{\pi}$ of dimension up to $(4N_d + 1)T$ if the W_{itjs}-policy is used or $(4N_d + 1)$ if the W_{it}-policy is used. Table 6 shows the number of variables required in the respective MIPs. Even with the W_i-policy or W_{it}-policy, the resulting MIP is too large to solve for moderate-sized power systems. For example, for a 2736-bus test case considered in Section 3.6, using the W_{it}-policy under this method would lead to more than 250 million $\boldsymbol{\pi}$-variables. We need a solution method that is more scalable.

3.4.2 Constraint Generation

Since the constraints in the affine multistage robust UC model have the form of (26), where the left-hand side is a linear function in $\boldsymbol{d}$ and the uncertainty set $\mathcal{D}$ is a

Table 6: Number of variables in the MIPs obtained directly using the duality-based approach.

Policy structure	W_i	W_{it}	W_{itjs}
Binary variables	$3N_gT$	$3N_gT$	$3N_gT$
$(\boldsymbol{w}, \boldsymbol{W})$-variables	$N_gT + N_g$	$2N_gT$	$N_gT + N_dN_gT(T+1)/2$
$\boldsymbol{\pi}$-variables	$8N_d(2N_g + N_l)T$	$8N_d(2N_g + N_l)T$	$4N_d(2N_g + N_l)T(T+1)$

polytope, each robust constraint is equivalent to an enumeration of the finitely many extreme points of the uncertainty set, in the following form:

$$\boldsymbol{c}^\top \boldsymbol{d} \leq h \qquad \forall\, \boldsymbol{d} \in \text{ext}(\mathcal{D}), \tag{28}$$

where $\text{ext}(\mathcal{D}) = \{d_1^*, ..., d_N^*\}$ is the set of extreme points of $\mathcal{D}$ (see [6]). This applies to every robust inequality in the affine multistage robust UC model. Furthermore, the energy balance equality constraints in the affine multistage robust UC model can be reformulated using the full-dimensionality of the uncertainty sets (equivalently, the existence of an interior point).

Proposition 5. *For a full-dimensional uncertainty set $\mathcal{D}$, the robust energy balance equation* (23g) *of the W_i-policy and W_{it}-policy is equivalent to the following equalities*

$$W_i\text{-Policy:} \quad \sum_{i \in \mathcal{N}_g} w_i^t = 0, \quad \sum_{i \in \mathcal{N}_g} W_i = 1 \qquad \forall t \in \mathcal{T} \tag{29}$$

$$W_{it}\text{-Policy:} \quad \sum_{i \in \mathcal{N}_g} w_i^t = 0, \quad \sum_{i \in \mathcal{N}_g} W_{it} = 1 \quad \forall t \in \mathcal{T}. \tag{30}$$

With the above observations, we can reformulate the multistage affine robust UC model (23) in the following compact form:

$$\min_{\boldsymbol{y} \in Y} \quad f(\boldsymbol{y}) \tag{31a}$$

$$\text{s.t.} \quad g_k(\boldsymbol{y}, \boldsymbol{d}) \leq 0 \qquad \forall \boldsymbol{d} \in \text{ext}(\mathcal{D}), \quad \forall k \in \{1, \ldots, K\}, \tag{31b}$$

where $\boldsymbol{y} = (\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}, \boldsymbol{W}, z)$ includes all decision variables in (23), the objective $f(\boldsymbol{y})$ represents (23a), and the set Y in (31a) is defined by (16b)-(16g) and (29) or (30)

according to the policy structure used. Constraints (31b) represent (23b)-(23f), where $g_k(\boldsymbol{y}, \boldsymbol{d})$ is a bilinear function in $\boldsymbol{y}$ and $\boldsymbol{d}$, and $K = 1 + 2T(2N_g + N_l)$ in (31b) is the total number of robust constraints.

This reformulation suggests a constraint generation framework. It starts with an initial set of extreme points for each robust constraint, and at each iteration, finds the worst-case scenario $\boldsymbol{d}$ for each robust constraint that achieves the highest constraint violation and adds it to the master problem, which is defined as

$$\begin{aligned} (MP) \quad & \min_{\boldsymbol{y} \in Y} \quad f(\boldsymbol{y}) \\ & \text{s.t.} \quad g_k(\boldsymbol{y}, \boldsymbol{d}) \leq 0 \quad \forall\, \boldsymbol{d} \in D_k, \quad \forall\, k \in \{1, ..., K\}, \end{aligned} \tag{32}$$

where $D_k \subseteq \text{ext}(\mathcal{D})$ is the list of extreme points that are identified from the constraint generation procedure for each robust constraint k in (31b). The constraint generation framework is outlined in Algorithm 2.

Algorithm 2 Constraint generation

1: Start with some initial D_k for all $k \in \{1, ..., K\}$
2: **repeat**
3: $\quad \boldsymbol{y}' \leftarrow$ optimal solution of the Master Problem (70).
4: $\quad$ **for all** $k \in \{1, ..., K\}$ **do**
5: $\quad\quad \boldsymbol{d}_k \leftarrow \text{argmax}_{d \in \mathcal{D}}\, g_k(\boldsymbol{y}', \boldsymbol{d})$
6: $\quad\quad$ If $g_k(\boldsymbol{y}', \boldsymbol{d}_k) > 0$ let $D_k \leftarrow D_k \cup \{\boldsymbol{d}_k\}$
7: $\quad$ **end for**
8: **until** $g_k(\boldsymbol{y}', \boldsymbol{d}_k) \leq 0$ for all $k \in \{1, ..., K\}$
9: **output**: $\boldsymbol{y}'$ is an optimal solution to (31)

Proposition 6. *The constraint generation algorithm presented in Algorithm 2 for solving the affine multistage robust UC problem* (23) *with uncertainty sets defined in* (18) *converges to the global optimum or reports infeasibility in a finite number of steps.*

Proof. Proof: The finite convergence follows from the fact that the uncertainty sets in (18) are bounded polyhedrons with a finite number of extreme points. □

3.4.3 Discussion

The constraint generation framework of Algorithm 1 can be also viewed as an embodiment of the cutting-plane method. A similar framework has been used in solving infinitely constrained optimization problems (see e.g. [18]). [35] studied the computational performance of a similar cutting-plane algorithm for solving static robust integer and linear programs with uncertainty in the constraint coefficients and using budgeted uncertainty sets. They find that the cutting-plane algorithm is more efficient than the duality-based approach for solving uncertain linear programs, and is less efficient when the problem involves integer decisions. Recently, [10] extended this comparison to different types of uncertainty sets including ellipsoidal uncertainty sets and explored different algorithmic strategies. Similar conclusions are reached. [11] presented a solution method based on constraint generation for adaptive optimization problems with a specific type of decision rule that can handle adaptive integer variables. Despite these interesting works, the computational study of solving large-scale multistage robust optimization problems with mixed-integer decisions still seems to be at an early stage.

In comparison to the above mentioned works, the affine multistage robust UC problem has some special characteristics such as high dimensionality in the numbers of continuous and integer variables and constraints, and also specific structures that can be exploited. Another point worth making is that the constraint generation framework of Algorithm 1 is equivalent to applying a Benders decomposition procedure to the MIP obtained with the duality-based approach, i.e., the feasibility cut generated through Benders decomposition on the dual system is equivalent to the primal constraint generated by constraint generation. As we will show in our experiments, the duality-based approach fails to solve even moderate sized problems with the simplest affine policy structure due to the scalability issue. Constraint generation, or the cutting-plane method, becomes necessary to handle this problem. However, a

direct implementation of Algorithm 1 also has limited success. Exploiting the special structure of the affine multistage robust UC model is crucial to devise an efficient constraint generation method.

3.5 Algorithmic Improvements

The constraint generation framework summarized in Algorithm 2 is still not efficient enough to handle large-scale problems. However, it does provide a basis for further algorithmic improvements, which proves to be critical in making the large-scale affine multistage robust UC model efficiently solvable. In the following, we develop an efficient procedure for the separation problem, an effective initialization for the master problem, a method to reduce the number of MIPs solved in the algorithm, and formulations to fully exploit the special structure of the W_i-policy and W_{it}-policy. We would also like to remark that the constraint generation framework with the proposed algorithmic improvements are not restricted to solving the robust UC problem, but can be applied to solve general multistage robust optimization problems with affine policies.

3.5.1 Efficient Separation Procedure

The separation procedure in the constraint generation algorithm involves solving the problem

$$\max_{\boldsymbol{d}\in\mathcal{D}} \; g_k(\boldsymbol{y}, \boldsymbol{d}) \tag{33}$$

for each robust constraint k in (23), in each iteration of the master problem. Thus, it is important to solve it as fast as possible. We can exploit two special structures of (33). First, as discussed above, $g_k(\boldsymbol{y}, \boldsymbol{d})$ is a linear function in $\boldsymbol{d}$ for any fixed $\boldsymbol{y}$. Second, the structure of the budgeted uncertainty set (18) allows us to solve the separation problem (33) by a simple sorting procedure, as we show below.

Proposition 7. *Consider the separation problem* $\max_{\boldsymbol{d}\in\mathcal{D}} \boldsymbol{c}^\top \boldsymbol{d}$*, where the uncertainty set* $\mathcal{D}$ *is defined in* (18)*. An optimal solution for this problem is given by* $(d_j^s)^* = \overline{d}_j^s + \Gamma \hat{d}_j^s (u_j^s)^*$ *for each* $s \in \mathcal{T}$*, where* $(u_j^s)^*$ *is obtained by the following procedure: let* $\{|c^s_{\sigma(j)}\, \hat{d}^s_{\sigma(j)}|\}_{j\in\mathcal{N}_d}$ *be a non-increasing ordering of* $\{|c_j^s\, \hat{d}_j^s|\}_{j\in\mathcal{N}_d}$*, where* $\sigma(\cdot)$ *determines the indices of the non-increasing order, and* $(u_j^s)^*$ *is given as follows:*

$$(u^s_{\sigma(j)})^* = \begin{cases} \operatorname{sign}(c^s_{\sigma(j)}) & \text{if } \sigma(j) \le \lfloor\sqrt{N_d}\rfloor, \\ (\sqrt{N_d} - \lfloor\sqrt{N_d}\rfloor)\cdot \operatorname{sign}(c^s_{\sigma(j)}) & \text{if } \sigma(j) = \lfloor\sqrt{N_d}\rfloor + 1, \\ 0 & \text{if } \sigma(j) \ge \lfloor\sqrt{N_d}\rfloor + 2, \end{cases}$$

where $\operatorname{sign}(x) = 1$ *if* $x \ge 0$ *and* -1 *otherwise.*

3.5.2 Initialization with Specific Uncertainty Scenarios

The constraint generation approach consists of iteratively finding extreme points of the uncertainty sets for each robust constraint until all robust constraints are satisfied. If there are extreme points that we believe to be strong candidates for being violated at some point in the constraint generation procedure, it would be useful to add them in the beginning.

Consider the vector $\boldsymbol{d}_{max}$ that achieves the maximum total net load in each time period, i.e., each component of $\boldsymbol{d}_{max}$ is defined as

$$\boldsymbol{d}^t_{max} \in \operatorname{argmax}_{\boldsymbol{d}\in\mathcal{D}} \sum_{j\in\mathcal{N}_d} d_j^t \quad \forall\, t \in \mathcal{T}. \tag{34}$$

This net load vector is clearly an important scenario in the uncertainty set for determining the worst-case dispatch costs. Thus, to speed up the constraint generation algorithm, we add $\boldsymbol{d}_{max}$ to D_k in the worst-case dispatch cost constraint (23b).

Similarly, we consider the minimum total net load $\boldsymbol{d}_{min} \in \mathcal{D}$ for (23b), which is defined as

$$\boldsymbol{d}^t_{min} \in \operatorname{argmin}_{\boldsymbol{d}\in\mathcal{D}} \sum_{j\in\mathcal{N}_d} d_j^t \quad \forall\, t \in \mathcal{T}. \tag{35}$$

We can also add $\boldsymbol{d}_{min}$ and $\boldsymbol{d}_{max}$ to D_k for every k representing the generation upper and lower bound constraints (23c).

For robust ramping constraints (23d) and (23e), consider the following scenarios for each t:

$$\boldsymbol{d}_{minmax}(t) = \left(\boldsymbol{d}^1_{min}, ..., \boldsymbol{d}^{t-1}_{min}, \boldsymbol{d}^t_{max}, ..., \boldsymbol{d}^T_{max}\right), \tag{36}$$

$$\boldsymbol{d}_{maxmin}(t) = \left(\boldsymbol{d}^1_{max}, ..., \boldsymbol{d}^{t-1}_{max}, \boldsymbol{d}^t_{min}, ..., \boldsymbol{d}^T_{min}\right), \tag{37}$$

which are the net loads with the largest up or down variations at period t. At initialization, we add $\boldsymbol{d}_{min}$, $\boldsymbol{d}_{max}$, $\boldsymbol{d}_{minmax}(t)$, $\boldsymbol{d}_{maxmin}(t)$ to D_k for every k representing the ramping constraints (23d)-(23e) at time t.

3.5.3 Complete Characterization for the W_{it}-Policy

The initialization technique in Section 3.5.2 is applicable to any affine policy. However, it has a very important consequence for the W_{it}-policy. Essentially, the robust constraints for generation limits and ramping can be completely characterized by a few uncertainty scenarios identified above, when using the W_{it}-policy. The computational benefit is huge.

Recall that the W_{it}-policy is described as $p^t_i(\boldsymbol{d}) = w^t_i + W_{it} \sum_{j \in \mathcal{N}_d} d^t_j$. When using the W_{it}-policy structure or any simpler policy such as the W_i-policy, generation output constraints (23c) and ramping constraints (23d)-(23e) are exactly equivalent to only considering the respective $\boldsymbol{d}$'s identified in (34)-(37), as we show below.

Proposition 8. *Under the W_{it}-policy or any simpler policy, and using the uncertainty set in* (18)*, the following statements hold:*

(i) The robust constraints on generation limits (23c) *are equivalent to the ones with the uncertainty set $\mathcal{D}$ replaced by the finite set $\{\boldsymbol{d}^t_{min}, \boldsymbol{d}^t_{max}\}$, where $\boldsymbol{d}_{min}$ and $\boldsymbol{d}_{max}$ are defined in* (35) *and* (34)*, respectively.*

(ii) The robust constraints on ramping limits (23d)-(23e) *at time t are equivalent to the ones with the uncertainty set $\mathcal{D}$ replaced by the finite set*

$$\{\boldsymbol{d}_{min}, \boldsymbol{d}_{max}, \boldsymbol{d}_{minmax}(t), \boldsymbol{d}_{maxmin}(t)\},$$

where $\boldsymbol{d}_{minmax}(t)$ and $\boldsymbol{d}_{maxmin}(t)$ are defined in (36) *and* (37)*, respectively.*

In the proof for ramping constraints we use the fact that the uncertainty set, given in (18), is separable over time periods.

This result implies that if we use the W_{it}-policy or any simpler policy such as the W_i-policy, the robust constraints corresponding to generation output limits and ramping capacities can be pre-computed before starting the constraint generation process. The only robust constraints left to deal with using constraint generation are the worst-case dispatch cost constraint (23b) and the transmission constraints (23f). This saves a tremendous amount of time checking feasibility and generating violated constraints. The overall convergence time of the constraint generation algorithm is significantly reduced.

3.5.4 Generating Multiple Cuts to the Master Problem

The difficulty in solving the affine multistage robust UC lies in finding all the necessary uncertainty scenarios $\boldsymbol{d}$'s for each robust constraint. This can lead to the undesired situation of solving the master problem (70) many times, which itself is a MIP with a large number of constraints. To strengthen the master problem, we employ a procedure that generates constraints by keeping all the binary variables fixed in the master problem. This can be helpful in reducing the number of MIP problems solved in the overall algorithm.

In particular, fix the commitment vector at the current solution $(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v})$ of the master problem, then the master problem becomes a linear program (LP) in the dispatch policy variables $(\boldsymbol{w}, \boldsymbol{W})$. Apply constraint generation to the resulting problem,

Algorithm 3 Generating multiple cuts for a fixed $\boldsymbol{x}'$

input: x', $\{D_k\}_{k=1}^K$
repeat
  $\boldsymbol{y}' = (\boldsymbol{x}', \boldsymbol{u}', \boldsymbol{v}', \boldsymbol{w}', \boldsymbol{W}', z') \leftarrow$ optimal solution of the master problem (70) with $\boldsymbol{x} = \boldsymbol{x}'$ fixed
  for all $k \in \{1, ..., K\}$ **do**
    $\boldsymbol{d}_k \leftarrow \text{argmax}_{\boldsymbol{d} \in \mathcal{D}}\ g_k(\boldsymbol{y}', \boldsymbol{d})$
    If $g_k(\boldsymbol{y}', \boldsymbol{d}_k) > 0$ let $D_k \leftarrow D_k \cup \{\boldsymbol{d}_k\}$
  end for
until $g_k(\boldsymbol{y}', \boldsymbol{d}_k) \leq 0$ for all $k \in \{1, ..., K\}$
output: $\{D_k\}_{k=1}^K$

starting from the current set of uncertainty scenarios $\boldsymbol{d}$'s until all the violated scenarios are identified for each robust constraint. This procedure is presented in Algorithm 3.

Furthermore, this technique can also be applied at the initialization phase of the overall constraint generation method. In particular, we can solve a static robust UC, which we define as a simplification of (23a)-(23g) by forcing $\boldsymbol{W} = \boldsymbol{0}$ and replacing robust energy balance constraints (23g) by enforcing it only for maximum total net load in the uncertainty set. This problem is very fast to solve and provides a good starting point for $\boldsymbol{x}$.

The concept of generating several cuts in each iteration of a constraint generation framework has been studied before with different formats. For example, [16] extended the L-shaped method for two-stage stochastic linear programs to a multicut version where each subproblem can induce a different cut. We make use of this idea in our algorithm, where each subproblem corresponds to checking the feasibility of a robust constraint. However, the enhancement presented here is different in that we proceed with the constraint generation algorithm solving an LP master problem with fixed binary variables, as many times as needed, inducing the fast generation of many "useful cuts" before solving each MIP master problem where binary variables are allowed to change. Another relevant idea that could be explored to enhance the algorithm is the concept of on-demand accuracy; see the work by [28] and references

therein. For example, some of the subproblems could be solved partially, and as the method develops the quality of their solutions could be increased as needed, potentially making the overall algorithm faster.

3.5.5 Algorithm Summary

The overall constraint generation algorithm with the above proposed algorithmic improvements is summarized in Algorithm 4. The initialization consists of finding $\boldsymbol{d}$'s described in Sections 3.5.2 and 3.5.3, and solving the static robust UC described in Section 3.5.4. Then the algorithm solves the master problem, and updates in each iteration the lists $\{D_k\}_{k=1}^K$ using each commitment solution found as described in Section 3.5.4.

Algorithm 4 Proposed solution method

$D_k \leftarrow \emptyset \quad \forall\, k = 1, 2, \ldots, K$
Add $\boldsymbol{d}$ from (34) to the D_k representing (23b)
Add $\boldsymbol{d}$'s from (34)-(35) to all D_k's representing (23c)
Add respective $\boldsymbol{d}$'s from (34)-(37) to all D_k's representing (23d)-(23e)
$\boldsymbol{x}' \leftarrow$ optimal solution of static robust UC
repeat
    Update $\{D_k\}_{k=1}^K$ using Algorithm 3 for $\boldsymbol{x}'$
    $\boldsymbol{y}' = (\boldsymbol{x}', \boldsymbol{u}', \boldsymbol{v}', \boldsymbol{w}', \boldsymbol{W}', z') \leftarrow$ optimal solution of (70)
    for all $k \in \{1, ..., K\}$ **do**
        $\boldsymbol{d}_k \leftarrow \text{argmax}_{\boldsymbol{d} \in \mathcal{D}}\ g_k(\boldsymbol{y}', \boldsymbol{d})$
        If $g_k(\boldsymbol{y}', \boldsymbol{d}_k) > 0$ let $D_k \leftarrow D_k \cup \{\boldsymbol{d}_k\}$
    end for
until $g_k(\boldsymbol{y}', \boldsymbol{d}_k) \leq 0$ for all $k \in \{1, ..., K\}$
output: $\boldsymbol{y}' = (\boldsymbol{x}', \boldsymbol{u}', \boldsymbol{v}', \boldsymbol{w}', \boldsymbol{W}', z')$ is an optimal solution for (31)

For simplicity, in our description of this algorithm we ignore the case where the master problem (70) reports infeasibility at some point. If such event ever occurs, the algorithm stops and reports infeasibility of the affine multistage robust UC problem under the affine policy used. Also, notice that checking for violated robust constraints can be parallelized, because it consists of solving K separate problems with the procedure described in Section 3.5.1.

3.6 Computational Experiments

We conduct extensive computational experiments on the IEEE 118-bus and the 2736-bus Polish systems (c.f. [120]). The major aspects of these instances are summarized in Table 7. In all cases, the UC problems involve a planning horizon of $T = 24$ hours. Uncertain net loads are located at every node with electricity demand. The uncertainty sets are given by (18), where we choose $\hat{d}_j^t = 0.1\overline{d}_j^t$ with various budget levels Γ, unless stated otherwise. All the experiments have been implemented using Python 2.7 in a PC laptop with an Intel Core i5 at 2.4 GHz and 4GB memory with CPLEX 12.5 as MIP and LP solver.

Table 7: Summary of test cases used

Buses	118	2736
Units	54	289
Loads	99	2011
Lines	186	100
Total generation capacity (MW)	7106	28880
Min total nominal net load (MW)	3327	10851
Max total nominal net load (MW)	4931	18075

Section 3.6.1 demonstrates the computational efficiency of the proposed algorithm. Section 3.6.2 shows that the simplified affine policies, as an approximation to the fully-adaptive policy, achieve close-to-optimal performance. Section 3.6.3 studies the impact of the UC solutions on the real-time dispatch operation from a worst-case perspective. In particular, it compares the worst-case performance of the real-time dispatch problem based on the UC solutions obtained from the two-stage robust UC model against those obtained from the affine multistage robust UC model. Section 3.6.4 studies the average performance of the affine multistage robust UC model in a rolling horizon simulation framework, and compares it with the deterministic and two-stage robust UC models.

3.6.1 Computational Performance of the Proposed Algorithm

In this section, we demonstrate the efficiency of the proposed solution methods for solving the affine multistage robust UC model in (23) with the W_{it}-policy structure. We show the efficiency enhancement achieved by individual algorithmic improvement techniques as well as the ultimate improvement achieved by their combination, and compare them with the two traditional solution methods, namely the duality-based approach (DBA) introduced in Section 3.4.1 and the basic constraint generation (CG) algorithm discussed in Section 3.4.2.

More specifically, we show the performance of the proposed algorithmic improvements in the following order. (a) The algorithm based on basic CG and Algorithm 3, which generates Multiple Cuts (MC) in each iteration for a fixed commitment solution (see Section 3.5.4). We denote this procedure as "CG + MC". (b) The algorithm based on basic CG and the method that exploits the Problem Structure (PS) of the W_{it}-policy (see Section 3.5.3). We denote this procedure as "CG + PS". (c) The combination of (a) and (b), denoted as "CG + MC + PS". (d) The combination of (a)(b)(c) along with the generation of an Initial Scenario (IS) of specific $\boldsymbol{d}$ for the worst-case dispatch cost constraint (see Section 3.5.2). This is the final solution algorithm summarized in Section 3.5.5. We denote it as "CG + MC + PS + IS".

All of the above four algorithms are implemented to solve the multistage robust UC model (23) with the W_{it}-policy on the 118-bus system. Table 8 shows the solution time (in seconds) of all these methods on the 118-bus system with different values of budget Γ for the uncertainty sets in (18). The stopping criterion of 0.1% optimality gap is used for solving each MIP problem. A time limit of 15,000 seconds is imposed on each algorithm. "M" and "T" in Table 8 stand for out-of-memory and out-of-time limits, respectively.

Notice that DBA and the basic CG are not efficient in solving the simple W_{it}-policy for the 118-bus system — either running out of memory or time limits. Applying the

Table 8: Solution time (seconds) of various algorithms for solving affine multistage robust UC under the W_{it}-policy for the 118-bus system.

Method	$\Gamma = 0.25$	$\Gamma = 0.5$	$\Gamma = 1$	$\Gamma = 2$	$\Gamma = 3$	$\Gamma = 4$
DBA	M	M	M	M	M	M
CG	T	T	T	T	T	T
CG + MC	6,807	8,475	5,639	3,488	10295	6,965
CG + PS	563	80	961	1,011	1183	1,227
CG + MC + PS	175	67	77	78	161	218
CG + MC + PS + IS	66	64	47	63	155	178

techniques of fixing the UC solution to find $\boldsymbol{d}$'s (CG + MC) or exploiting the policy structure (CG + PS) leads to a substantial improvement in solution times, especially when the special structure of the W_{it}-policy is exploited (CG + PS). When the two techniques are combined (i.e., CG + MC + PS), the solution times are reduced to within 218 seconds (less than 4 minutes) for all sizes of tested uncertainty sets, and even faster for problems with small uncertainty sets (around 1 minute). Running time is further reduced by initializing the algorithm with one more valid $\boldsymbol{d}$ for the worst-case dispatch cost constraint (CG + MC + PS + IS).

At this point, let us try to understand why the algorithm with "CG + MC + PS + IS" is significantly more efficient than DBA. For the medium-sized 118-bus system with the W_{it}-policy, the MIP obtained by directly applying DBA requires the creation of approximately 70 million dual variables and 35 million associated constraints for the explicit representation of all the robust constraints of the original formulation. If the special structure of the W_{it}-policy is exploited (which can also be combined with DBA) these numbers are reduced to approximately 3.5 million dual variables and 1.8 million associated constraints, which still runs into memory issues. In contrast, the algorithm with "CG + MC + PS + IS" requires initially creating about 15,000 constraints for policy structure exploitation, but does not require the creation of dual variables. Furthermore, if we take $\Gamma = 4$, a total number of 263 constraints are generated in the master problem along the algorithm, where a total of 5 MIPs and 7

LPs are solved (while these numbers are even smaller for the other Γ's tested). This significantly saves the computation time comparing to a naive CG method.

Algorithm "CG + MC + PS + IS", thus identified as the most efficient algorithm among the six tested methods, is applied to the 2736-bus Polish system. Table 9 presents the solution times of this algorithm for solving the multistage robust UC model with the W_{it}-policy structure, for different values of Γ. For the 2736-bus system, an optimality gap of 1% is used for all MIP problems solved in the algorithm. In Table 9, "inf" indicates that the algorithm detects the problem being infeasible, which is caused by the large size of the uncertainty set. The solution time variations for the 2736-bus system are explained by the variability in the time taken for the MIPs to be solved.

Table 9: Solution time using "CG + MC + PS + IS" algorithm for both systems studied under the W_{it}-policy.

System	$\Gamma = 0.25$	$\Gamma = 0.5$	$\Gamma = 1$	$\Gamma = 2$	$\Gamma = 3$	$\Gamma = 4$
118-bus	66s	64s	47s	63s	155s	178s
2736-bus	3.6h	3.2h	2.3h	2.0h	2.4h	0.4h (inf)

For the 2736-bus Polish system, when $\Gamma = 1$, a total number of 6 MIPs and 5 LPs are solved, and 727 constraints are generated by the proposed constraint generation algorithm. Similar numbers are obtained for the other values of Γ tested. In comparison, if DBA is used with exploitation of the special structure of the W_{it}-policy, the MIP obtained would require approximately 39 million dual variables and 19 million associated constraints for the explicit representation of all the robust constraints of the original formulation. Furthermore, without exploiting the structure of the W_{it}-policy, more than 250 million dual variables would be required in DBA.

From Table 9 we can see that the proposed algorithm can efficiently solve the real-world 2736-bus system within a time framework reasonable for the day-ahead operation. Considering the complexity of the multistage robust UC model and the simple computation resources (a moderate personal computer) that our experiments

rely on, these computational experiments show that the affine multistage robust UC model and the proposed algorithms are very promising for practical applications in large-scale power system operations.

3.6.2 Optimality Gap for Simplified Affine Policies

The affine multistage robust UC model proposed in (23) is an approximation scheme to the original fully-adaptive multistage robust UC model (20). The UC solution and the affine dispatch policy thus obtained are feasible, but may not be optimal for the fully-adaptive model. In this section, we study the approximation quality of the simplified affine policies. As will be shown, affine policies with the very simple W_i-policy or W_{it}-policy perform surprisingly well as approximate solutions to the fully adaptive problem. This is a particularly encouraging result for the large-scale 2736-bus system.

3.6.2.1 Bounding the Approximation Quality of Affine Policies

The two-stage robust UC formulation (17) is a relaxation of the fully adaptive multistage robust UC model (20) by ignoring non-anticipativity on dispatch decisions. Thus, the optimal objective value of the two-stage robust UC problem, denoted as v^*_{2S}, provides a *lower* bound to the optimal objective value of the fully adaptive multistage robust UC, denoted as v^*_{MS}. However, obtaining a globally optimal solution of the two-stage robust UC problem for large-scale power systems is still computationally challenging (e.g. see [13]). To reduce computation time, we employ the heuristic used by [67], which generates a *lower* bound to v^*_{2S}, denoted as $\underline{v}_{2S}$. Furthermore, since the affine policy is an approximation to the fully adaptive policy, its optimal objective value, denoted as v^*_{AFF}, provides an *upper* bound to the optimal objective value of the fully adaptive multistage robust UC. Because the MIP solver is terminated within a certain accuracy (e.g. with a 0.1% MIP gap), the solution

at termination gives a further upper bound to v^*_{AFF}, denoted as $\overline{v}_{AFF}$. In summary, we have the following relations between objective values of different solutions: $\underline{v}_{2S} \leq v^*_{2S} \leq v^*_{MS} \leq v^*_{AFF} \leq \overline{v}_{AFF}$. Then, the optimality gap between v^*_{AFF} and v^*_{MS}, i.e., $(v^*_{AFF} - v^*_{MS})/v^*_{MS}$, is upper bounded as

$$0 \leq \frac{v^*_{AFF} - v^*_{MS}}{v^*_{MS}} \leq \frac{\overline{v}_{AFF} - \underline{v}_{2S}}{\underline{v}_{2S}} \triangleq \text{Guaranteed Optimality Gap.}$$

We call the upper bound to the optimality gap the *guaranteed optimality gap* of the affine multistage robust UC problem.

3.6.2.2 Computational Results for Guaranteed Optimality Gap

Table 10 presents the guaranteed optimality gaps of two simple affine policy structures for the 118-bus and 2736-bus systems with different values of the uncertainty set size parameter Γ.

Table 10: Guaranteed opt. gap under different policy structures ("inf" indicates infeasibility).

118-bus system

Policy	$\Gamma = 0.25$	$\Gamma = 0.5$	$\Gamma = 1$	$\Gamma = 1.5$	$\Gamma = 2$	$\Gamma = 3$	$\Gamma = 4$
W_i	0.04%	0.02%	0.04%	0.08%	0.10%	0.26%	0.67%
W_{it}	0.04%	0.02%	0.03%	0.07%	0.07%	0.17%	0.35%

2736-bus system

Policy	$\Gamma = 0.25$	$\Gamma = 0.5$	$\Gamma = 1$	$\Gamma = 1.5$	$\Gamma = 2$	$\Gamma = 3$	$\Gamma = 4$
W_i	0.09%	0.22%	0.42%	0.55%	1.05%	inf	inf
W_{it}	0.07%	0.11%	0.25%	0.35%	0.53%	0.94%	inf

From these results, we offer the following observations.

1. For each test system, the W_{it}-policy achieves a better guaranteed optimality gap than the W_i policy, especially for larger uncertainty sets. For example, for the 2736-bus system with $\Gamma = 2$, the guaranteed optimality gap is improved from 1.05% of the W_i-policy to 0.53% by the W_{it}-policy. For smaller uncertainty sets, the W_i-policy has a more comparable performance to the W_{it}-policy.

2. The simple W_{it}-policy achieves surprisingly good performance in both test cases. The guaranteed optimality gap is at most 0.94% for all sizes of uncertainty sets in both test systems. Due to its strong performance and computational tractability, we will use the W_{it}-policy in all the following experiments.

3.6.3 Worst-Case Performance Analysis

As discussed in Section 3.2.3, the proposed multistage robust UC formulation is motivated by a critical issue of the two-stage robust UC model, namely that it ignores non-anticipativity in the dispatch process for the sequential revelation of uncertain net loads, and thus may not be prepared in real-time operations for all realizations of net loads within the uncertainty set. Indeed, Claims 1 and 2 in Section 3.2.3 show that this is possible based on a simple two-bus example. This section will further study this issue on the 118-bus and the 2736-bus systems. In particular, we want to estimate "how much" infeasibility can be caused in the real-time dispatch under the commitment solutions of the two-stage robust UC model. For this purpose, the two-stage model is solved for different sizes of the uncertainty sets, then the obtained UC solutions are fed into the affine multistage robust model (23). That is, the UC decision in (23) is fixed at the two-stage UC solution, and the remaining affine multistage robust dispatch problem is solved. The dispatch model is properly augmented with penalty variables in the energy balance and transmission constraints, so that the degree of infeasibility can be quantified by the amount of penalty costs incurred (see Section 3.9 in Appendix 3.9 for details on the penalty variables). In this way, we can compare the worst-case operational costs (including penalty costs) of the real-time dispatch under the two-stage robust UC solutions against those obtained under the affine multistage robust UC solutions. It is important to carry out this type of worst-case performance study of the real-time dispatch under different UC solutions, because power system operations require extremely high reliability. Infeasibility in

real-time operation has to be resolved by starting expensive fast-start units or shedding load, both of which bear significant economic consequences.

Table 11 presents the results. "Total Cost" is the worst-case dispatch cost plus penalty cost of the affine multistage robust dispatch model under a specific UC solution. "Penalty" is the total penalty cost associated with constraint violations in the dispatch model, where \$5000/MW is used as the unit penalty cost. "Rel Diff" is the relative difference between the total costs obtained by the multistage and two-stage UC solutions. We can make the following observations.

1. The multistage UC solutions do not cause any infeasibility in real-time operation for $\Gamma \leq 3$, whereas even though the two-stage UC model is feasible in itself for both 118-bus and 2736-bus systems, its UC solutions cause infeasibility to the multistage robust dispatch and incur quite significant penalties in the real-time operation.

2. The penalty costs and the total costs of the two-stage UC solutions increase as the size of the uncertainty set grows. For the 118-bus system, the two-stage model has 62.87% more total cost than the multistage model at $\Gamma = 3$, and the penalty cost is over \$1.2M. For the 2736-bus system, the two-stage UC model incurs 25.70% more total cost than the multistage model at $\Gamma = 3$, and the absolute amount of penalty cost exceeds \$2.7M.

These results further demonstrate the importance of non-anticipative constraints and the multistage robust UC model in power system operations.

We have the following further discussion. In the above experiment, the same uncertainty set sizes Γ are used in the two-stage and multistage UC models. It is also interesting to test if the two-stage UC solution would perform better in the multistage dispatch if the two-stage UC model uses a larger value of Γ than the Γ later used in the multistage dispatch problem. In this way, a larger Γ might "compensate" the

Table 11: Worst-case cost (US$) of multistage robust dispatch under the two-stage and multistage UC solutions. Multistage models use the W_{it}-policy.

118-bus system

	$\Gamma = 0.5$	$\Gamma = 1$	$\Gamma = 1.5$	$\Gamma = 2$	$\Gamma = 3$
Affine multistage UC solutions					
Total Cost	1,696,304	1,725,470	1,755,398	1,784,543	1,845,218
Penalty	0	0	0	0	0
Two-stage UC solutions					
Total Cost	1,696,456	1,749,766	1,797,503	1,897,212	3,005,290
Penalty	0	52,501	55,268	196,101	1,229,300
Rel Diff	0.01%	1.41%	2.40%	6.31%	62.87%

2736-bus system

	$\Gamma = 0.5$	$\Gamma = 1$	$\Gamma = 1.5$	$\Gamma = 2$	$\Gamma = 3$
Affine multistage UC solutions					
Total Cost	9,445,069	9,596,788	9,746,685	9,905,527	10,234,459
Penalty	0	0	0	0	0
Two-stage UC solutions					
Total Cost	9,505,651	9,745,889	10,183,433	10,975,403	12,864,719
Penalty	96,313	224,952	591,661	1,165,324	2,703,522
Rel Diff	0.64%	1.55%	4.49%	10.80%	25.70%

two-stage UC solution for its lack of non-anticipativity. For this purpose, we feed the two-stage UC solutions obtained using $\Gamma = 3$ to the multistage dispatch problem with $\Gamma = 0.5, 1, 1.5, 2$. For the 2736-bus system, the worst-case costs of the total UC costs thus obtained are respectively 1.35%, 1.04%, 0.84% and 0.66% higher than those obtained by directly solving the affine multistage robust UC problem, which indeed are better than the performance reported in Table 11. From this we can see that if the two-stage robust UC is solved under a conservative "over-robustness" request, better solutions can be obtained than using smaller Γ's, however the performance is still not as cost-effective as those obtained by directly solving the multistage robust UC problem. Similar results are obtained for the 118-bus system.

3.6.4 Average Performance of UC Models in Real-Time Dispatch

In the previous section, we have conducted a worst-case analysis to compare the two-stage and multistage robust UC models. In this section, we study the average performance of different UC solutions and their impact on real-time dispatch. We develop a rolling-horizon simulation platform to mimic the real time operation of the power system, where information about uncertain net load is revealed sequentially as time moves forward. On this platform, we conduct Monte-Carlo simulations of different economic dispatch (ED) models that are suitable for the associated UC solution concepts. In particular, we propose a new robust ED model that exploits the affine policy obtained from the multistage robust UC model. For the two-stage robust UC and the deterministic UC solutions, we use a multi-period ("look-ahead") deterministic ED model in simulation, which has started to be adopted in some ISO markets (the most prevalent in practice is still the single-period ED model) ([34, Table 4], [76]).

3.6.4.1 Efficient Robust Dispatch Model Exploiting Affine Policy.

The proposed robust ED model is motivated by the following considerations. First, solving the affine multistage robust UC model not only produces a UC solution, but also provides an affine policy that could be exploited in the ED process. Second, any ED model needs to be solved fast within a few minutes in real-time operation.

With these considerations, we propose a new robust dispatch model in (38), which we call the *policy-enforcement* robust ED model. At each time t, the dispatch decision $\boldsymbol{p}^t$ is the first-stage decision, which satisfies all the dispatch constraints $\Omega_t(\boldsymbol{x}, \boldsymbol{d}^t, \boldsymbol{p}^{t-1})$ in the current period and will be implemented "right now" at time t. Furthermore, the policy-enforcement robust ED model also considers the next period's dispatch decision $\boldsymbol{p}^{t+1}$ and assumes that it takes the form of the affine policy with coefficients $(w_i^{t+1}, \boldsymbol{W}_i^{t+1})$ of time $t+1$ obtained from the day-ahead affine multistage robust UC

model.

$$\min_{\boldsymbol{p}^t} \quad \sum_{i\in\mathcal{N}_g} C_i p_i^t \tag{38a}$$

$$\text{s.t.} \quad \boldsymbol{p}^t \in \Omega_t(\boldsymbol{x}, \boldsymbol{d}^t, \boldsymbol{p}^{t-1}) \tag{38b}$$

$$w_i^{t+1} + \boldsymbol{W}_i^{t+1}\boldsymbol{d}^{t+1} - p_i^t \geq -RD_i x_i^{t+1} - SD_i v_i^{t+1} \quad \forall \boldsymbol{d}^{t+1} \in \mathcal{D}^{t+1} \tag{38c}$$

$$w_i^{t+1} + \boldsymbol{W}_i^{t+1}\boldsymbol{d}^{t+1} - p_i^t \leq RU_i x_i^t + SU_i u_i^{t+1} \quad \forall \boldsymbol{d}^{t+1} \in \mathcal{D}^{t+1}. \tag{38d}$$

Here, $\Omega_t(\boldsymbol{x}, \boldsymbol{d}^t, \boldsymbol{p}^{t-1})$ includes all the dispatch related constraints in the deterministic UC model (16) at time t, with the observed values of the current period's net load vector $\boldsymbol{d}^t$ and the previous period's dispatch level $\boldsymbol{p}^{t-1}$. Constraints (38d) and (38c) enforce ramping limits between $\boldsymbol{p}^t$ and $\boldsymbol{p}^{t+1}$ for any realization of nodal net loads in the uncertainty set at time $t+1$. In this way, the proposed dispatch model coordinates the ramping capabilities in the two consecutive periods and hedges against unfavorable net load realizations in future periods.

It is important to note that we can also consider a multi-period model where affine policies obtained from the multistage robust UC model for all future periods $t+1, t+2, \ldots$ are used. However, this multi-period model is exactly equivalent to the above two-period model, because the affine policies obtained from the robust UC model already satisfy all the dispatch constraints in each future period as well as the ramping constraints coupling every two consecutive periods. Also notice that the above robust ED model has almost the same size as a deterministic single-period ED, since we can use the strategy in Section 3.5.3 to handle robust ramping constraints.

For the deterministic and two-stage robust UC solutions, there is no affine policy readily available to exploit. Instead, we use the deterministic multi-period look-ahead ED model for their dispatch simulation, where net loads in future periods use forecast values (i.e., the nominal $\overline{d}_j^t$ values), and the ED model is obtained from the deterministic UC model (16) by fixing the commitment decision.

3.6.4.2 Rolling-Horizon Simulation Platform for Real-Time Dispatch.

We develop a rolling horizon platform to simulate the real-time dispatch process. In particular, for each UC solution, we select an ED model according to the discussion in Section 3.6.4.1. At each time period t in the simulation, the selected ED model is solved with the observation of nodal net load up to time t, and the dispatch solution of time period t is implemented. Then the time horizon rolls forward and the same procedure is repeated. This simulation process is different from the existing ones in the literature such as in [13, 53, 118], where net loads over the entire scheduling horizon are revealed all at once to the dispatch model, ignoring non-anticipativity.

We consider a 24-hour horizon with an hourly step size in the simulation process. At each time t, the robust ED model in (38) considers two periods t and $t+1$, i.e., a one period look-ahead, whereas the deterministic look-ahead ED model considers 4 periods, i.e. a three periods look-ahead. The look-ahead horizon shrinks in the last three periods. Each round of the rolling-horizon simulation contains $T = 24$ consecutive runs of the ED model through the entire horizon. For each UC solution and the corresponding ED model, we carry out multiple rounds of such simulations. In the uncertainty set of the robust UC problems solved, we used $\overline{d}_j^t$ and $\hat{d}_j^t$ selected equal to the expected value and standard deviation, respectively, of net load at bus j and time t. The same set of nodal net load trajectories are used in all evaluations of different UC solutions to generate a fair comparison. Penalty variables are incorporated to deal with violations of energy balance and transmission, all of which have a unit penalty cost of \$5000/MWh. Due to space restriction, we only show the results for the 2736-bus system.

3.6.4.3 Results for the 2736-bus system with temporally independent nodal net loads.

In the experiments presented in this subsection, the nodal net load corresponds to demand sampled from a normal distribution with a standard deviation equal to 10%

of its expected value (i.e., $\hat{d}_j^t = 0.10\bar{d}_j^t$) and independent across time periods. In order for the uncertainty set (18) to contain the entire ℓ_∞-ball of $\Pi_{j\in\mathcal{N}_d}[\bar{d}_j^t - \hat{d}_j^t, \bar{d}_j^t + \hat{d}_j^t]$, the budget parameter Γ has to be $\sqrt{N_d} = \sqrt{2011} = 44.8$. This corresponds to an extremely conservative robust solution. We want to choose a Γ value that is small enough to still guarantee a robust enough performance. The following experiments use $\Gamma \leq 3.0$, which corresponds to significantly smaller uncertainty sets and less conservative solutions. In each experiment, 1000 rounds of simulation are conducted.

Table 12 presents the simulation performance of the multistage robust UC solution with the W_{it}-policy and the corresponding policy-enforcement robust ED model, the two-stage robust UC solution with the deterministic look-ahead ED model, and a deterministic reserve approach, for the 2736-bus system. We compare the average total costs over the 24-hour horizon ("Cost Avg"), their standard deviation ("Cost Std"), the average penalty costs ("Penalty Avg"), and the average frequency of penalty occurrence ("Penalty Freq Avg"). We also study the performance of the deterministic UC model with adjusted reserve and look-ahead ED in the rolling-horizon simulation, which resembles the current operational practice. The reserve adjustment follows the rule used in [13] with various reserve levels tested.

From these results we can see that the multistage robust UC model achieves the best average total cost at $\Gamma = 0.5$, which is a 0.46% $((9319396 - 9362379)/9362379)$ reduction from the best average cost of the two-stage robust UC model achieved at $\Gamma = 3$, and a 0.95% reduction from that of the deterministic UC with reserve adjusted at 20%. Further comparing these three columns, we can see that the multistage robust UC solution achieves a significant improvement on system reliability, with a cost standard deviation reduced by 64.97% from the two-stage solution and 82.69% from the deterministic UC with reserve. Moreover, the penalty cost of the multistage robust UC solution is reduced by 42.70% and 98.43% from the two-stage robust UC and deterministic UC solutions, respectively. The penalty cost can be reduced to zero

Table 12: Simulation performance of the different models for the 2736-bus system with temporally independent demand.

Affine multistage robust UC with policy-enforcement robust ED

Γ	0.25	0.5	1	1.5	2	3
Cost Avg ($)	9,397,528	9,319,396	9,342,754	9,360,359	9,379,464	9,442,858
Cost Std ($)	113,725	15,970	12,828	12,509	12,363	12,092
Penalty Cost Avg ($)	93,552	3497	727	61	5	0
Penalty Freq Avg	10.00%	1.47%	0.40%	0.01%	0.00%	0.00%

Two-stage robust UC with look-ahead ED

Γ	0.25	0.5	1	1.5	2	3
Cost Avg ($)	9,398,109	9,456,599	9,408,732	9,383,569	9,407,290	9,362,379
Cost Std ($)	93,470	195,774	173,884	144,698	162,469	45,584
Penalty Cost Avg ($)	80,127	152,637	98,113	66,801	82,864	6,103
Penalty Freq Avg	9.93%	12.26%	7.80%	5.11%	5.57%	0.37%

Deterministic UC with reserve and look-ahead ED

Reserve	2.5%	5%	10%	15%	20%	30%
Cost Avg ($)	9,556,549	9,575,446	9,424,678	9,561,024	9,408,173	9,411,741
Cost Std ($)	261,464	288,777	121,122	196,354	92,268	69,050
Penalty Cost Avg ($)	254,627	271,672	119,127	248,658	83,938	51,907
Penalty Freq Avg	15.93%	13.37%	14.31%	18.16%	10.03%	7.22%

by a larger value of Γ in the multistage model, whereas both the two-stage robust UC and deterministic UC do not achieve zero penalty for all tested budget and reserve levels.

3.6.4.4 Results for the 2736-bus system with persistent demand.

Here we present simulation results for the 2736-bus system where nodal net corresponds to demand sampled from a persistent model [45], which exhibits some simple temporal correlation. In particular, we sample the trajectory of demand at bus j from the following autoregressive model:

$$\widetilde{d}_j^t = \mu_j^t + \sigma_j^t Z_j^t \quad \forall\, t \in \mathcal{T}$$

$$Z_j^t = \phi Z_j^{t-1} + \epsilon_j^t \quad \forall\, t \in \mathcal{T},$$

where the ϵ_j^t's are sampled independently from a normal distribution with an expected value of 0 and a standard deviation of $\sigma^\epsilon = \sqrt{1-\phi^2}$, and Z_j^0 is sampled from a normal distribution with an expected value of 0 and a standard deviation of 1. The value of σ^ϵ is selected so that the standard deviation of Z_j^t is always 1. Further, $\sigma_j^t = 0.1\mu_j^t$.

Given this, $\widetilde{d}_j^t$ has an expected value of μ_j^t and a standard deviation of $0.1\mu_j^t$, and the correlation between $\widetilde{d}_j^t$ and $\widetilde{d}_j^{t-1}$ is ϕ. In the uncertainty set, we take $\overline{d}_j^t = \mu_j^t$ and $\hat{d}_j^t = \sigma_j^t = 0.1\mu_j^t$. The setting in Section 3.6.4.3 corresponds to the case $\phi = 0$. In this section we use $\phi = 0.9$ denoting a strong temporal correlation in demand.

Table 13 presents the simulation performance results for the different UC approaches studied. We can observe that the best average cost is still achieved at $\Gamma = 0.5$ for the multistage robust UC, at $\Gamma = 3$ for the two-stage robust UC, and at 20% reserve for the deterministic UC. For these cases, the multistage robust UC presents a 0.39% and 0.84% reduction in average cost with respect to the two-stage robust UC and the deterministic UC, respectively. We can also observe that with a proper choice for Γ the multistage robust UC can completely eliminate the penalty cost, while the two-stage robust UC and deterministic UC could not completely eliminate penalty under the tested values of Γ and reserve.

Table 13: Simulation performance of the different models for the 2736-bus system with persistent demand.

Affine multistage robust UC with policy-enforcement robust ED

Γ	0.25	0.5	1	1.5	2	3
Cost Avg ($)	9,395,199	9,320,462	9,343,907	9,361,336	9,380,499	9,443,846
Cost Std ($)	180,122	51,226	42,344	41,229	41,295	40,387
Penalty Cost Avg ($)	90,292	3,464	835	8	0	0
Penalty Freq Avg	10.11%	1.23%	0.43%	0.01%	0.00%	0.00%

Two-stage robust UC with look-ahead ED

Γ	0.25	0.5	1	1.5	2	3
Cost Avg ($)	9,390,163	9,409,159	9,380,783	9,386,101	9,362,994	9,357,111
Cost Std ($)	118,856	234,627	187,981	164,032	130,666	43,809
Penalty Cost Avg ($)	71,238	10,3861	69,188	60,483	45,171	5
Penalty Freq Avg	9.35%	11.68%	7.05%	5.30%	5.23%	0.01%

Deterministic UC with reserve and look-ahead ED

Reserve	2.5%	5%	10%	15%	20%	30%
Cost Avg ($)	9,525,854	9,565,603	9,415,143	9,515,113	9,398,984	9,409,452
Cost Std ($)	372,593	369,799	163,575	271,478	131,586	108,410
Penalty Cost Avg ($)	222,972	261,016	108,863	201,491	73,849	48,801
Penalty Freq Avg	15.05%	13.08%	13.58%	16.45%	8.83%	6.11%

3.6.4.5 Results for the 2736-bus system using temporally correlated wind data.

In the above simulations we sampled demand from specific distributions, first assuming temporal independence and then incorporating temporal correlations through an autoregressive model. However, in power systems with a high penetration of wind power, most of the uncertainty in net loads stems from the intermittency of wind power outputs, and we would like to study the performance of our approach under more realistic net load trajectories for such power systems. For this purpose, we added 140 wind farms to the 2736-bus system, with each wind farm located at a different load bus, making use of one year of wind power output data from 140 locations of NREL's Western Wind Integration Dataset [77, 85]. The average hourly total wind power output is 1051MW, which corresponds to 7.37% of average hourly total demand. Each of the $N = 365$ days of wind power output data corresponds to one simulation, and for each of these 24-hour trajectories we also consider demand generated as described in Section 3.6.4.2, namely from a normal distribution at each bus with a standard deviation $\sigma_{demand,tj}$ corresponding to 10% of the expected value at the respective bus j and time t. Then, we use linear regression to estimate daily and semi-daily seasonality pattern and subtract it from the wind power data to estimate the standard deviation $\sigma_{wind,tj}$ of the errors of wind farm j and time t. When building the uncertainty set for robust UC, we select $\overline{d}_j^t$ as the nominal value for net load at bus j and time t, calculated as the expected demand minus the expected wind power output, and $\hat{d}_j^t$ as the standard deviation of net load, calculated as $\sqrt{\sigma^2_{demand,tj} + \sigma^2_{wind,tj}}$, assuming independence between demand and wind power output.

The simulation performance results are presented in Table 14. We can see that now the multistage robust UC model achieves the best average total cost at $\Gamma = 1$, which is a 1.14% reduction from the best average cost of the two-stage robust UC model, achieved at $\Gamma = 2$. The standard deviation of the cost is reduced by 6.07%. Furthermore, we can see that the multistage robust UC approach can completely

remove penalty cost with $\Gamma = 3$, while penalties remain fairly large for the two-stage robust UC approach under all Γ's (we also tested Γ's larger than 3, confirming our statement). Comparing with the deterministic reserve-based approach, we can see that the deterministic UC becomes very ineffective at handling this level of temporally correlated uncertainty. In particular, the multistage robust UC model reduces the average cost by 24.52% from the deterministic UC, and reduces the standard deviation of the cost by 83.22%.

Table 14: Simulation performance of the different models for the 2736-bus system with wind power and temporally independent demand.

Affine multistage robust UC with policy-enforcement robust ED

Γ	0.25	0.5	1	1.5	2	3
Cost Avg ($)	11,078,125	9,511,651	8,523,553	8,568,135	8,642,679	9,414,582
Cost Std ($)	3,726,779	2,006,668	528,697	452,158	423,938	454,504
Penalty Cost Avg ($)	2,761,688	1,163,142	122,558	69,472	24,334	0
Penalty Freq Avg	18.80%	14.21%	1.95%	0.48%	0.15%	0.00%

Two-stage robust UC with look-ahead ED

Γ	0.25	0.5	1	1.5	2	3
Cost Avg ($)	10,431,937	11,368,530	8,780,404	8,904,157	8,622,251	8,944,327
Cost Std ($)	1,865,595	1,065,188	689,745	867,257	562,878	754,036
Penalty Cost Avg ($)	2,106,438	1,035,071	426,520	530,980	209,218	440,250
Penalty Freq Avg	13.29%	3.74%	7.84%	5.68%	2.15%	2.68%

Deterministic UC with reserve and look-ahead ED

Reserve	2.5%	5%	10%	15%	20%	30%
Cost Avg ($)	13,343,698	14,395,889	13,259,879	13,647,325	11,986,311	11,292,887
Cost Std ($)	5,652,913	7,032,034	5,634,975	6,025,708	4,134,035	3,150,160
Penalty Cost Avg ($)	5,064,315	6,128,731	4,979,369	5,365,449	3,686,461	2,957,919
Penalty Freq Avg	30.57%	29.99%	33.49%	32.85%	24.41%	15.59%

Finally, Table 15 presents simulation performance results under the temporally correlated wind power data described above, and with demand sampled from the persistent model in Section 3.6.4.4. The best average cost is still achieved at $\Gamma = 1$ for the multistage robust UC, at $\Gamma = 2$ for the two-stage robust UC, and at 30% reserve for the deterministic UC, with the multistage robust UC achieving a 1.23% and 24.52% reduction in average cost, with respect to the two-stage robust UC and deterministic UC, respectively. We can also observe that under $\Gamma = 3$ the multistage

robust UC does not completely eliminate the penalty cost, but achieves a significant reduction as compared to the other models, with a penalty frequency average of 0.01% as compared to 2.07% for the two-stage robust UC under $\Gamma = 2$ and 15.03% for the deterministic UC under 30% reserve (which respectively achieve their minimum penalties). All of these experiments on a large-scale power system demonstrate that the multistage UC model together with the proposed robust ED approach can dominate the performance of the two-stage robust UC and the deterministic UC models with look-ahead ED in terms of both average cost and system reliability.

Table 15: Simulation performance of the different models for the 2736-bus system with wind power and persistent demand.

Affine multistage robust UC with policy-enforcement robust ED

Γ	0.25	0.5	1	1.5	2	3
Cost Avg ($)	10,996,931	9,459,785	8,502,923	8,581,532	8,646,665	9,415,693
Cost Std ($)	3,665,301	2,007,317	490,457	466,999	424,801	458,865
Penalty Cost Avg ($)	2,679,299	1,110,032	101,234	81,834	27,344	218
Penalty Freq Avg	18.84%	14.44%	1.67%	0.47%	0.18%	0.01%

Two-stage robust UC with look-ahead ED

Γ	0.25	0.5	1	1.5	2	3
Cost Avg ($)	10,390,214	11,365,568	8,734,840	8,863,975	8,609,160	8,947,959
Cost Std ($)	1,831,279	1,059,427	620,301	802,441	522,881	793,447
Penalty Cost Avg ($)	2,064,045	1,032,109	380,451	490,562	195,681	443,401
Penalty Freq Avg	12.73%	3.68%	7.37%	5.19%	2.07%	2.66%

Deterministic UC with reserve and look-ahead ED

Reserve	2.5%	5%	10%	15%	20%	30%
Cost Avg ($)	13,186,705	14,272,477	13,110,030	13,617,194	11,879,817	11,248,546
Cost Std ($)	5,557,309	7,023,964	5,596,039	6,082,173	4,095,780	3,113,902
Penalty Cost Avg ($)	4,905,635	6,003,861	4,827,766	5,334,746	3,578,986	2,912,186
Penalty Freq Avg	30.45%	29.94%	33.00%	32.43%	23.61%	15.03%

3.7 Conclusion

This Chapter presented a systematic study of multistage adaptive robust optimization for the UC problem with the solution concept of simplified affine policy. Such a model can deal with significant uncertainty in electricity demand and renewable generation caused by a high level penetration of wind and solar resources. We also propose a solution framework based on constraint generation with various algorithmic improvements, which achieves efficient solution of the affine multistage robust

UC in large-scale power systems when the traditional methods fail. We also propose an associated robust ED model for real-time dispatch, which exploits the solution of the affine multistage robust UC model and is quickly solvable in real-time operation. We conduct extensive computational experiments on medium and large-scale power systems to thoroughly study the performance of the proposed models and algorithms and to compare them with existing approaches. The results show that the proposed algorithms can effectively solve the multistage robust UC model with simplified affine policies within a time frame reasonable for the day-ahead operation of large-scale power systems. The computational results demonstrate the effectiveness of the multistage robust UC model in reducing operational costs and at the same time improving system reliability, compared to the existing two-stage robust UC model and a deterministic UC model with reserve. Built on this work, future research can further explore more complex affine or non-affine policy structures and respective solution algorithms, and also modeling techniques to combine uncertainty from heterogeneous sources such as wind and solar power, demand, and generation or transmission contingencies.

3.8 Appendix: Proofs for Sections 3.3, 3.4, 3.5

Proof. Proof of Proposition 3: To make references more explicit, we use $(2S)$ and (M) to denote the two-stage (17) and the multistage models (20) in this proof, respectively. The proof follows from the fact that, without ramping constraints (20c), the dispatch problems using uncertainty sets (18) in both $(2S)$ and (M) are separable over time periods. In fact, we show that, without ramping constraints, $(2S)$ and (M) are both equivalent to problem $(1P)$, defined as follows:

$$(1P) \quad \min_{(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v})\in X} \Big\{ \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}_g} \big(G_i x_i^t + S_i u_i^t\big) + \sum_{t\in\mathcal{T}} \max_{\boldsymbol{d}^t\in\mathcal{D}^t} \min_{\boldsymbol{p}^t\in\Omega_t^{NR}(\boldsymbol{x},\boldsymbol{d}^t)} \sum_{i\in\mathcal{N}_g} C_i\, p_i^t \Big\},$$

where $X = \{(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}) : \text{ (16b)-(16g) are satisfied}\}$ and $\Omega_t^{NR}(\boldsymbol{x}, \boldsymbol{d}^t)$ is the feasible dispatch set at time t without ramping constraints, i.e.

$$\Omega_t^{NR}(\boldsymbol{x}, \boldsymbol{d}^t) \triangleq \{\boldsymbol{p}^t : \text{ (16h), (16j), (16k) are satisfied}\}.$$

(i) First, we show that without ramping constraints, $(2S)$ is equivalent to $(1P)$. In fact, without ramping constraints, $(2S)$ can be written as

$$\min_{(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v})\in X} \Big\{ \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}_g} \big(G_i x_i^t + S_i u_i^t\big) + \max_{\boldsymbol{d}\in\mathcal{D}} \min_{\{\boldsymbol{p}:\, \boldsymbol{p}^t\in\Omega_t^{NR}(\boldsymbol{x},\boldsymbol{d}^t)\,\forall t\in\mathcal{T}\}} \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}_g} C_i\, p_i^t \Big\},$$

and we have

$$\begin{aligned} & \max_{\boldsymbol{d}\in\mathcal{D}} \min_{\{\boldsymbol{p}:\, \boldsymbol{p}^t\in\Omega_t^{NR}(\boldsymbol{x},\boldsymbol{d}^t)\,\forall t\in\mathcal{T}\}} \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}_g} C_i\, p_i^t \\ = \; & \max_{\boldsymbol{d}\in\mathcal{D}} \sum_{t\in\mathcal{T}} \min_{\boldsymbol{p}^t\in\Omega_t^{NR}(\boldsymbol{x},\boldsymbol{d}^t)} \sum_{i\in\mathcal{N}_g} C_i\, p_i^t \\ = \; & \sum_{t\in\mathcal{T}} \max_{\boldsymbol{d}^t\in\mathcal{D}^t} \min_{\boldsymbol{p}^t\in\Omega_t^{NR}(\boldsymbol{x},\boldsymbol{d}^t)} \sum_{i\in\mathcal{N}_g} C_i\, p_i^t, \end{aligned}$$

where the first equality comes from the fact that the dispatch set $\{\boldsymbol{p} : \boldsymbol{p}^t \in \Omega_t^{NR}(\boldsymbol{x}, \boldsymbol{d}^t) \;\forall t \in \mathcal{T}\}$ is separable over time, and the second equality comes from the separability of the uncertainty set $\mathcal{D}$ defined in (18) over time periods. Adding $\sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}_g} (G_i x_i^t + S_i u_i^t)$ and applying $\min_{(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v})\in X}$ at both sides of this equality yields the desired result.

(ii) Now we show that, without ramping constraints, (M) is equivalent to $(1P)$. Without ramping constraints, $\Omega_t(\boldsymbol{x}, \boldsymbol{d}^t, \boldsymbol{p}^{t-1}) = \Omega_t^{NR}(\boldsymbol{x}, \boldsymbol{d}^t)$, so the nested multistage formulation (21) is equivalent to

$$\begin{aligned} (\widetilde{M}^{NR}) \quad \min_{(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v})\in X} \Big\{ & \boldsymbol{G}^\top \boldsymbol{x} + \boldsymbol{S}^\top \boldsymbol{u} \\ & + \max_{\boldsymbol{d}^1\in\mathcal{D}^1} \min_{\boldsymbol{p}^1\in\Omega_1^{NR}(\boldsymbol{x},\boldsymbol{d}^1)} \Big\{ \boldsymbol{C}^\top \boldsymbol{p}^1 + \cdots + \max_{\boldsymbol{d}^T\in\mathcal{D}^T} \min_{\boldsymbol{p}^T\in\Omega_T^{NR}(\boldsymbol{x},\boldsymbol{d}^T)} \boldsymbol{C}^\top \boldsymbol{p}^T \Big\} \Big\}. \end{aligned}$$

Consider the max-min problem at $t = T-1$ in $(\widetilde{M}^{NR})$. Since $\mathcal{D}^T$, $\Omega_T^{NR}(\boldsymbol{x}, \boldsymbol{d}^T)$, and $\boldsymbol{C}^\top \boldsymbol{p}^T$ do not depend on $\boldsymbol{p}^{T-1}$ and $\boldsymbol{d}^{T-1}$, we obtain

$$\max_{\boldsymbol{d}^{T-1}\in\mathcal{D}^{T-1}} \min_{\boldsymbol{p}^{T-1}\in\Omega_{T-1}^{NR}(\boldsymbol{x},\boldsymbol{d}^{T-1})} \Big\{ \boldsymbol{C}^\top \boldsymbol{p}^{T-1} + \max_{\boldsymbol{d}^T\in\mathcal{D}^T} \min_{\boldsymbol{p}^T\in\Omega_T^{NR}(\boldsymbol{x},\boldsymbol{d}^T)} \boldsymbol{C}^\top \boldsymbol{p}^T \Big\}$$

$$= \max_{\boldsymbol{d}^{T-1}\in\mathcal{D}^{T-1}} \left\{ \left(\min_{\boldsymbol{p}^{T-1}\in\Omega_{T-1}^{NR}(\boldsymbol{x},\boldsymbol{d}^{T-1})} \boldsymbol{C}^\top \boldsymbol{p}^{T-1} \right) + \left(\max_{\boldsymbol{d}^T\in\mathcal{D}^T} \min_{\boldsymbol{p}^T\in\Omega_T^{NR}(\boldsymbol{x},\boldsymbol{d}^T)} \boldsymbol{C}^\top \boldsymbol{p}^T \right) \right\}$$

$$= \left(\max_{\boldsymbol{d}^{T-1}\in\mathcal{D}^{T-1}} \min_{\boldsymbol{p}^{T-1}\in\Omega_{T-1}^{NR}(\boldsymbol{x},\boldsymbol{d}^{T-1})} \boldsymbol{C}^\top \boldsymbol{p}^{T-1} \right) + \left(\max_{\boldsymbol{d}^T\in\mathcal{D}^T} \min_{\boldsymbol{p}^T\in\Omega_T^{NR}(\boldsymbol{x},\boldsymbol{d}^T)} \boldsymbol{C}^\top \boldsymbol{p}^T \right),$$

and this argument can be carried out backward until $t = 1$ to obtain

$$\max_{\boldsymbol{d}^1\in\mathcal{D}^1} \min_{\boldsymbol{p}^1\in\Omega_1^{NR}(\boldsymbol{x},\boldsymbol{d}^1)} \left\{ \boldsymbol{C}^\top \boldsymbol{p}^1 + \cdots + \max_{\boldsymbol{d}^T\in\mathcal{D}^T} \min_{\boldsymbol{p}^T\in\Omega_T^{NR}(\boldsymbol{x},\boldsymbol{d}^T)} \boldsymbol{C}^\top \boldsymbol{p}^T \right\}$$

$$= \sum_{t\in\mathcal{T}} \max_{\boldsymbol{d}^t\in\mathcal{D}^t} \min_{\boldsymbol{p}^t\in\Omega_t^{NR}(\boldsymbol{x},\boldsymbol{d}^t)} \boldsymbol{C}^\top \boldsymbol{p}^t.$$

Adding $\sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}_g} (G_i x_i^t + S_i u_i^t)$ and applying $\min_{(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v})\in X}$ on both sides of this equality yields that $(\widetilde{M}^{NR})$ is equivalent to $(1P)$, which completes the proof.

□

Proof. Proof of Proposition 4: $\boldsymbol{c}^\top \boldsymbol{d} \le h \;\forall\, \boldsymbol{d} \in \mathcal{D}$ is equivalent to $\max_{\boldsymbol{d}\in\mathcal{D}} \boldsymbol{c}^\top \boldsymbol{d} \le h$. Now notice that $\mathcal{D}^t$ is the projection over d^t of

$$\widetilde{\mathcal{D}}^t = \left\{ (\boldsymbol{d}^t, \boldsymbol{z}^t) : \sum_{j\in\mathcal{N}_d} z_j^t \le \Gamma\sqrt{N_d}, \right.$$

$$\left. \hat{d}_j^t z_j^t \ge d_j^t - \overline{d}_j^t, \;\; \hat{d}_j^t z_j^t \ge \overline{d}_j^t - d_j^t, \;\; d_j^t \in [\overline{d}_j^t - \Gamma\hat{d}_j^t, \overline{d}_j^t + \Gamma\hat{d}_j^t] \;\; \forall\, j \in \mathcal{N}_d \right\}.$$

So by defining $\widetilde{\mathcal{D}} = \prod_{t\in[1:T]} \widetilde{\mathcal{D}}^t$, we have $\max_{\boldsymbol{d}\in\mathcal{D}} \boldsymbol{c}^\top \boldsymbol{d} = \max_{(\boldsymbol{d},\boldsymbol{z})\in\widetilde{\mathcal{D}}} \boldsymbol{c}^\top \boldsymbol{d} = \min_{\boldsymbol{\pi}\in\Pi} \boldsymbol{e}^\top \boldsymbol{\pi}$, where the last equality follows from duality theory, since $\mathcal{D}$ is bounded, where $\boldsymbol{\pi} \in \Pi$ is equivalent to (27b)-(27d) and $\boldsymbol{e}^\top \boldsymbol{\pi}$ is the left hand side of (27a). Now, $\min_{\boldsymbol{\pi}\in\Pi} \boldsymbol{e}^\top \boldsymbol{\pi} \le h$ is equivalent to the existence of $\boldsymbol{\pi} \in \Pi$ such that $\boldsymbol{e}^\top \boldsymbol{\pi} \le h$ and the result follows. □

Proof. Proof of Proposition 5: Take the W_{it}-policy. The energy balance equation (23g) can be written as

$$\sum_{i\in\mathcal{N}_g} w_i^t + \sum_{j\in\mathcal{N}_d} \left(\sum_{i\in\mathcal{N}_g} W_{it} - 1 \right) d_j^t = 0 \qquad \forall \boldsymbol{d} \in \mathcal{D}, \;\; \forall t \in \mathcal{T}.$$

Since the uncertainty set $\mathcal{D}$ is full-dimensional, which is the case for the uncertainty set in (18), the constraint that the above affine function of $\boldsymbol{d}$ is equal to zero for all $\boldsymbol{d} \in \mathcal{D}$ can hold if and only if all the coefficients of this affine function are zero, which gives (30). We can show (29) similarly. □

Proof. Proof of Proposition 7: In the separation problem, consider the change of variables given by $d_j^s = \overline{d}_j^s + \Gamma \hat{d}_j^s u_j^s$. The equivalent problem for $\boldsymbol{u}$ is

$$\begin{aligned} \max_{\boldsymbol{u}} \quad & \sum_{s \in \mathcal{T}} \sum_{j \in \mathcal{N}_d} c_j^s \, \hat{d}_j^s \, u_j^s \\ \text{s.t.} \quad & u_j^s \in [-1,\, 1] \quad \forall\, j \in \mathcal{N}_d,\, s \in \mathcal{T} \\ & \sum_{j \in \mathcal{N}_d} |u_j^s| \leq \sqrt{N_d} \quad \forall\, s \in \mathcal{T}. \end{aligned}$$

This problem is separable in s and the solution of each of the problems obtained is found by ordering $|c_j^s \, \hat{d}_j^s|$ in j from largest to smallest, and successively assigning the highest possible values to those $|u_j^s|$ with the largest respective values of $|c_j^s \, \hat{d}_j^s|$, taking each of these u_j^s with the same sign of c_j^s (notice that $\hat{d}_j^s > 0$). □

Proof. Proof of Proposition 8: The proof follows from reformulating the respective robust constraints.

(i) Constraints (23c) under the W_{it}-policy can be written, for each i and t, as

$$p_i^{min} x_i^t \leq w_i^t + W_{it} \left(\sum_{j \in \mathcal{N}_d} d_j^t \right) \leq p_i^{max} x_i^t \quad \forall \boldsymbol{d} \in \mathcal{D},$$

which is equivalent to

$$p_i^{min} x_i^t \leq w_i^t + \min_{\boldsymbol{d} \in \mathcal{D}} W_{it} \left(\sum_{j \in \mathcal{N}_d} d_j^t \right)$$

$$w_i^t + \max_{\boldsymbol{d}\in\mathcal{D}} W_{it}\left(\sum_{j\in\mathcal{N}_d} d_j^t\right) \le p_i^{max} x_i^t.$$

Depending on the sign of W_{it}, the above two inequalities are equivalent to the following four constraints,

$$p_i^{min} x_i^t \le w_i^t + W_{it} \min_{\boldsymbol{d}\in\mathcal{D}}\left(\sum_{j\in\mathcal{N}_d} d_j^t\right) \le p_i^{max} x_i^t$$
$$p_i^{min} x_i^t \le w_i^t + W_{it} \max_{\boldsymbol{d}\in\mathcal{D}}\left(\sum_{j\in\mathcal{N}_d} d_j^t\right) \le p_i^{max} x_i^t.$$

In other words, in the robust constraints (23c) $\mathcal{D}$ can be replaced by a finite uncertainty set consisting of $\{\boldsymbol{d}_{min}, \boldsymbol{d}_{max}\}$. This completes the proof for the first part. Notice that the proof of this part is independent of the structure of $\mathcal{D}$ so the conclusion of (i) is true for any convex uncertainty set.

(ii) Under the W_{it}-policy, (23e) can be written as

$$\max_{\boldsymbol{d}\in\mathcal{D}} \left\{ W_{it}\left(\sum_{j\in\mathcal{N}_d} d_j^t\right) - W_{i,t-1}\left(\sum_{j\in\mathcal{N}_d} d_j^{t-1}\right)\right\} \le w_i^{t-1} - w_i^t + RU_i x_i^{t-1} + SU_i u_i^t. \tag{39}$$

Notice that the uncertainty set $\mathcal{D}$ defined in (18) is separable in time periods. Therefore, the left-hand side of (39) is equivalent to the following problem

$$\max_{\boldsymbol{d}^t\in\mathcal{D}^t} \left\{ W_{it}\left(\sum_{j\in\mathcal{N}_d} d_j^t\right)\right\} - \min_{\boldsymbol{d}^{t-1}\in\mathcal{D}^{t-1}} \left\{ W_{i,t-1}\left(\sum_{j\in\mathcal{N}_d} d_j^{t-1}\right)\right\}. \tag{40}$$

Depending on the signs of the affine coefficients W_{it} and $W_{i,t-1}$, (40) is equivalent to one of the four possible combinations:

$$W_{it} \max_{\boldsymbol{d}^t\in\mathcal{D}^t} \left(\sum_{j\in\mathcal{N}_d} d_j^t\right) - W_{i,t-1} \min_{\boldsymbol{d}^{t-1}\in\mathcal{D}^{t-1}} \left(\sum_{j\in\mathcal{N}_d} d_j^{t-1}\right) \tag{41a}$$
$$W_{it} \max_{\boldsymbol{d}^t\in\mathcal{D}^t} \left(\sum_{j\in\mathcal{N}_d} d_j^t\right) - W_{i,t-1} \max_{\boldsymbol{d}^{t-1}\in\mathcal{D}^{t-1}} \left(\sum_{j\in\mathcal{N}_d} d_j^{t-1}\right) \tag{41b}$$
$$W_{it} \min_{\boldsymbol{d}^t\in\mathcal{D}^t} \left(\sum_{j\in\mathcal{N}_d} d_j^t\right) - W_{i,t-1} \min_{\boldsymbol{d}^{t-1}\in\mathcal{D}^{t-1}} \left(\sum_{j\in\mathcal{N}_d} d_j^{t-1}\right) \tag{41c}$$

$$W_{it} \min_{\boldsymbol{d}^t \in \mathcal{D}^t} \left(\sum_{j \in \mathcal{N}_d} d_j^t \right) - W_{i,t-1} \max_{\boldsymbol{d}^{t-1} \in \mathcal{D}^{t-1}} \left(\sum_{j \in \mathcal{N}_d} d_j^{t-1} \right), \tag{41d}$$

where (41a)-(41d) correspond to the worst-case scenarios $\boldsymbol{d}_{maxmin}(t)$, $\boldsymbol{d}_{max}$, $\boldsymbol{d}_{min}$, $\boldsymbol{d}_{minmax}(t)$, respectively. The proof is analogous for ramping down constraints (23d).

□

3.9 Appendix: Incorporating Penalty Variables

In order to incorporate penalty variables into the affine multistage robust UC with W_{it}-policy we replace equations (23b), (23f) and (23g) by

$$\sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{N}_g} C_i \left(w_i^t + W_{it} \sum_{j \in \mathcal{N}_d} d_j^t \right)$$

$$+ C^{pen} \sum_{t \in \mathcal{T}} \left[\left(w_t^+ + W_t^+ \sum_{j \in \mathcal{N}_d} d_j^t \right) + \left(w_t^- + W_t^- \sum_{j \in \mathcal{N}_d} d_j^t \right) + \sum_{l \in \mathcal{N}_l} w_{tl}^f \right] \leq z \quad \forall \, \boldsymbol{d} \in \mathcal{D}$$

$$- f_l^{max} - w_{tl}^f \leq \sum_m \sum_{i \in \mathcal{N}_g} \alpha_{lm} B_{mi}^p \left(w_i^t + W_{it} \sum_{j \in \mathcal{N}_d} d_j^t \right) - \sum_m \sum_{j \in \mathcal{N}_d} \alpha_{lm} B_{mj}^d d_j^t \leq f_l^{max} + w_{tl}^f$$

$$\forall \, \boldsymbol{d} \in \mathcal{D}, \, t \in \mathcal{T}, \, l \in \mathcal{N}_l$$

$$\sum_{i \in \mathcal{N}_g} \left(w_i^t + W_{it} \sum_{j \in \mathcal{N}_d} d_j^t \right) + \left(w_t^+ + W_t^+ \sum_{j \in \mathcal{N}_d} d_j^t \right) = \sum_{j \in \mathcal{N}_d} d_j^t + \left(w_t^- + W_t^- \sum_{j \in \mathcal{N}_d} d_j^t \right)$$

$$\forall \boldsymbol{d} \in \mathcal{D}, t \in \mathcal{T}$$

$$w_{tl}^f \geq 0 \quad \forall \, t \in \mathcal{T}, \, l \in \mathcal{N}_l$$

$$\left(w_t^+ + W_t^+ \sum_{j \in \mathcal{N}_d} d_j^t \right), \, \left(w_t^- + W_t^- \sum_{j \in \mathcal{N}_d} d_j^t \right) \geq 0 \quad \forall \, \boldsymbol{d} \in \mathcal{D}, \, t \in \mathcal{T},$$

where C^{pen} is the unitary penalty cost, the w_{tl}^f's are penalty variables for transmission line capacity constraints, and the w_{it}^+, W_{it}^+, w_{it}^- and W_{it}^-'s are the penalty variables for under- and over-generation in the energy balance equation.

CHAPTER IV

MULTISTAGE ROBUST UNIT COMMITMENT WITH DYNAMIC UNCERTAINTY SETS AND ENERGY STORAGE

4.1 Introduction

The reliable and cost-effective operation of power systems with an abundant presence of wind and solar power depends critically on the competence of optimization methods to effectively manage their uncertainty. The most crucial decision process that faces this challenge is the unit commitment (UC) problem, which schedules generating capacities for the next day and prepares the power system for potentially strong variations in the availability of intermittent renewable resources.

Many efforts within the realm of optimization under uncertainty have been developed for the UC problem. Two main types of methods are stochastic programming and robust optimization. Typically, stochastic programming methods involve scenario trees for modeling uncertain parameters, see for example [25], [83], [97], [101], and references therein. The stochastic programming framework is versatile, however, it may induce substantial computational difficulties for large-scale problems, and it is difficult to properly represent temporal and spatial correlations within scenario trees. Robust optimization instead relies on the concept of uncertainty set, namely, a deterministic set of realizations of uncertain parameters, which leads to simplified models and improved computational tractability. The research on robust UC is growing rapidly, starting with a static robust model for the contingency constrained UC proposed in [94] and the two-stage robust UC models first developed in [13], [53], and [118]. In [13] a security constrained robust UC model is developed under net load uncertainty,

[53] presents a robust UC model with wind uncertainty and pumped-storage units, and [118] develops a robust UC model with wind uncertainty and demand response. Other two-stage robust UC approaches have considered generator and transmission line contingencies [103], the combination with stochastic UC [115], a worst-case regret objective [54], and the use of exact and heuristic approaches to solve bilinear subproblems [55], [96], to name a few works. In two-stage robust UC models, generator on/off decisions are selected in the first stage and then dispatch decisions are made in the second stage with the full knowledge of all the uncertain parameters in the future. This assumption on the knowledge of future uncertainty in the two-stage model is unrealistic, because in the real-life dispatch process the operators only know realizations of uncertainty up to the time of the dispatch decision.

A more accurate UC model should restrict operators' actions to only depend on uncertain parameters realized up to the current decision period. That is, the *non-anticipativity* of dispatch decisions needs to be enforced. Such a robust model is called a *multistage* robust UC model. The benefit of such a model over two-stage robust and deterministic UC models is that the multistage model properly prepares the system's ramping capability to meet future demand variations, which is especially important for systems with limited ramping capacity and significant renewable variations [69].

Multistage robust optimization in its most general form is computationally intractable, but the concept of affine policy has been proposed as an effective approximation, where recourse decisions take the form of an affine function of uncertain parameters [7]. In order to handle the large scale of the UC problem, simplified affine policies were proposed in [69], as well as a solution method based on constraint generation. Another application of affine policies for UC can be found in a stochastic UC model in [107], where the classic approach of duality reformulation is used and computational results are shown for a 2-dimensional uncertainty set and a 118-bus

system. In the context of power systems, the affine policy approach has also been applied in a stochastic economic dispatch problem with energy storage decisions [105], in a robust optimal power flow problem [51], and in a chance-constrained optimal power flow problem [15].

A key component of any robust optimization model is the uncertainty set used to represent uncertain parameters [29]. A particularly important challenge is to capture the correlational structure of uncertain parameters. Most of the existing literature, including the above robust optimization references, have considered *static* uncertainty sets where temporal and spatial correlations are not systematically represented or simply ignored. However, some important efforts have been undertaken to try to improve these uncertainty models. In [22], primitive uncertainties with potential asymmetric distributions are considered as underlying factors that determine the uncertain parameters of interest, which can be used to capture dependencies. In [72], state-space representable uncertainty sets are considered, which can also be used to capture certain dependencies. In [67], the idea of dynamic uncertainty sets is proposed to capture temporal and spatial correlations in wind speeds.

In this Chapter, we improve the modeling and solution methods for multistage robust UC problems in several important directions. In particular, this Chapter addresses the following research questions: (i) How to incorporate temporal and spatial dynamics of uncertain wind and solar outputs in the multistage robust UC model? (ii) How to efficiently solve the resulting multistage robust UC models for large-scale systems with high dimensional uncertainty? (iii) How to effectively utilize the dispatch policy obtained from the multistage robust UC model in real-time dispatch? (iv) What is the impact of using energy storage in the multistage robust UC model? The contributions of this Chapter are summarized as follows:

1. We propose a new multistage robust UC model with both wind and solar power uncertainty and energy storage, using a simple and effective affine policy. In

this model wind and solar power are dispatchable and their availability is the uncertain component.

2. We also formulate a new robust look-ahead economic dispatch model that utilizes the affine dispatch policy obtained from the multistage robust UC to improve the robustness of real-time operation.

3. We develop a data-driven approach to construct dynamic uncertainty sets for capturing joint temporal and spatial correlations of multiple wind and solar farms, including a critical enhancement to reduce the dimensionality of these sets. These sets directly model renewable power, rather than rely on explanatory factors such as wind speed or solar irradiance.

4. We develop an efficient solution method combining constraint generation and the duality based approach with various algorithmic enhancements, including the use of outer approximation techniques for reformulating inter-temporal constraints, the use of a one-tree Benders implementation, and constraint screening speed-up techniques. The proposed algorithm can solve multistage robust UC models for the Polish 2736-bus power system with high-dimensional uncertainty within a couple of hours on a modest personal computer, which offers a practical solution for real-world operations.

5. Extensive computational experiments on a simulation platform that mimics the hour-to-hour operation of a real-world power system are carried out to show the benefits of the proposed models and algorithms in comparison to other robust and deterministic approaches.

The remainder of this Chapter is organized as follows. Section 4.2 formulates the multistage robust UC model and the dispatch policy, as well as an economic dispatch (ED) method that exploits the robust adaptive dispatch policy. Section 4.3 proposes

dynamic uncertainty sets for modeling wind and solar power. Section 4.4 develops the solution algorithms. Section 4.5 presents computational experiments. Section 4.6 concludes.

4.2 Multistage Robust Unit Commitment Model

4.2.1 Fully-Adaptive Model

We describe here the fully-adaptive multistage robust UC model considered in this Chapter. The decisions in this model are composed of binary on/off commitment decisions of conventional generators, dispatch decisions of conventional generators and renewable units, and charge/discharge decisions of storage units. The uncertainty considered in this model corresponds to the power availability at renewable units. The purpose of the problem is to determine on/off decisions that are decided before the uncertainty is revealed, and a dispatch policy that determines how all other decisions adapt to uncertainty as it is revealed throughout time, with the objective of minimizing the total worst-case cost. The notation is as follows. Let $\mathcal{N}^d, \mathcal{N}^g, \mathcal{N}^l, \mathcal{N}^r, \mathcal{N}^s, \mathcal{T}$ denote the sets of demand nodes, generators, transmission lines, renewable units (wind or solar), storage units, and time periods, respectively. For the different types of units the corresponding notation is described below.

1. For *conventional* generator i at time t: x^o_{it}, x^+_{it} and x^-_{it} are the on/off, start-up, and shut-down decisions, jointly denoted by vector $\boldsymbol{x}$; $p^g_{it}(\overline{\boldsymbol{p}}^r_{[t]})$ is the adaptive dispatch *policy* of generator output, which is a function of the uncertain available renewable power realized up to time t; $\boldsymbol{c}$ is a vector encompassing no-load, start-up, and shut-down costs of generators, and C^g_i is the variable cost of generator i; $\underline{p}^g_{it}$ and $\overline{p}^g_{it}$ are the minimum and maximum output levels of generator i; RD_{it} and RU_{it} are the ramp-down and ramp-up rates of generator i, and SD_{it} and SU_{it} are the ramp rates when generator i shuts down or turns on, respectively.

2. For *renewable* unit i at time t: $\overline{p}^r_{it}$ denotes its available power output, which is an uncertain parameter in the robust UC model (42); $\overline{\mathcal{P}}^r$ is the uncertainty set for the uncertain vector $\overline{\boldsymbol{p}}^r$ of available renewable output; $p^r_{it}(\cdot)$ is the dispatch policy of renewable unit i.

3. For *storage* unit i at time t: $p^{s+}_{it}(\cdot)$ and $p^{s-}_{it}(\cdot)$ are respectively its discharge and charge policies; $\underline{p}^{s+}_{it}$, $\overline{p}^{s+}_{it}$, $\underline{p}^{s-}_{it}$, and $\overline{p}^{s-}_{it}$ are the limits for its power output and input; and q^s_{i0}, $\overline{q}^s_i$, and η^s_i are its initial storage level, capacity, and efficiency.

Based on the description above, vector $\boldsymbol{x}$ encompasses all binary decisions related to generator commitment, and $\boldsymbol{p}(\cdot) = (\boldsymbol{p}^g(\cdot), \boldsymbol{p}^r(\cdot), \boldsymbol{p}^s(\cdot))$ corresponds to a dispatch policy of conventional, renewable, and storage units that are functions of the available renewable power $\overline{\boldsymbol{p}}^r$. For the transmission system, let α^d_{lj}, α^g_{li}, α^r_{li}, and α^s_{li} denote the corresponding shift factor values for demands, generators, renewable units, and storage units, respectively, for transmission line l, based on a DC power flow model, and let f^{max}_l denote the flow limit for line l. Finally, let d^t_j denote the demand level of node j at time t.

With the above notation, the fully adaptive robust UC model is formulated as follows:

$$\min_{\boldsymbol{x},\boldsymbol{p}(\cdot)} \left\{ \boldsymbol{c}^\top x \; + \; \max_{\overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r} \sum_{i \in \mathcal{N}_g} \sum_{t \in \mathcal{T}} C^g_i \, p^g_{it}(\overline{\boldsymbol{p}}^r_{[t]}) \right\} \tag{42a}$$

$$\text{s.t.} \;\; \boldsymbol{x} \in X \tag{42b}$$

$$x^o_{it} \, \underline{p}^g_{it} \le p^g_{it}(\overline{\boldsymbol{p}}^r_{[t]}) \le x^o_{it} \, \overline{p}^g_{it} \quad \forall \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r, i \in \mathcal{N}^g, t \in \mathcal{T} \tag{42c}$$

$$- RD_{it} \, x^o_{it} - SD_{it} \, x^-_{it} \le p^g_{it}(\overline{\boldsymbol{p}}^r_{[t]}) - p^g_{i,t-1}(\overline{\boldsymbol{p}}^r_{[t-1]}) \le RU_{it} \, x^o_{i,t-1} + SU_{it} \, x^+_{it} \tag{42d}$$

$$\forall \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r, i \in \mathcal{N}^g, t \in \mathcal{T}$$

$$0 \le p^r_{it}(\overline{\boldsymbol{p}}^r_{[t]}) \le \overline{p}^r_{it} \quad \forall \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r, i \in \mathcal{N}^r, \, t \in \mathcal{T} \tag{42e}$$

$$\underline{p}^{s+}_{it} \le p^{s+}_{it}(\overline{\boldsymbol{p}}^r_{[t]}) \le \overline{p}^{s+}_{it} \quad \forall \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r, i \in \mathcal{N}^s, \, t \in \mathcal{T} \tag{42f}$$

$$\underline{p}^{s-}_{it} \le p^{s-}_{it}(\overline{\boldsymbol{p}}^r_{[t]}) \le \overline{p}^{s-}_{it} \quad \forall \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r, i \in \mathcal{N}^s, \, t \in \mathcal{T} \tag{42g}$$

$$0 \leq q^s_{i0} + \sum_{\tau \in [1:t]} \left(\eta^s_i \, p^{s-}_{i\tau}(\overline{\boldsymbol{p}}^r_{[\tau]}) - p^{s+}_{i\tau}(\overline{\boldsymbol{p}}^r_{[\tau]}) \right) \leq \overline{q}^s_i \quad \forall \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r,\, i \in \mathcal{N}^s,\, t \in \mathcal{T} \tag{42h}$$

$$- f^{max}_l \leq \sum_{i \in \mathcal{N}^g} \alpha^g_{li} \, p^g_{it}(\overline{\boldsymbol{p}}^r_{[t]}) + \sum_{i \in \mathcal{N}^r} \alpha^r_{li} \, p^r_{it}(\overline{\boldsymbol{p}}^r_{[t]}) + \sum_{i \in \mathcal{N}^s} \alpha^s_{li} \, \left(p^{s+}_{it}(\overline{\boldsymbol{p}}^r_{[t]}) - p^{s-}_{it}(\overline{\boldsymbol{p}}^r_{[t]}) \right) \tag{42i}$$

$$- \sum_{j \in \mathcal{N}^d} \alpha^d_{lj} \, d_{jt} \leq f^{max}_l \quad \forall \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r,\, l \in \mathcal{N}^l,\, t \in \mathcal{T}$$

$$\sum_{j \in \mathcal{N}^d} d_{jt} = \sum_{i \in \mathcal{N}^g} p^g_{it}(\overline{\boldsymbol{p}}^r_{[t]}) + \sum_{i \in \mathcal{N}^r} p^r_{it}(\overline{\boldsymbol{p}}^r_{[t]}) + \sum_{i \in \mathcal{N}^s} \left(p^{s+}_{it}(\overline{\boldsymbol{p}}^r_{[t]}) - p^{s-}_{it}(\overline{\boldsymbol{p}}^r_{[t]}) \right) \tag{42j}$$

$$\forall \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r,\, t \in \mathcal{T}.$$

In this problem, the objective (42a) consists of minimizing the sum of commitment costs (including no-load, start-up, and shut-down costs) and the worst-case dispatch cost, which is assumed to be linear for ease of exposition, but can be replaced with a piecewise linear function without changing the structure of the problem. Notice that the second-stage decision in (42) is a policy function. The overall $\min - \max$ structure in (42a) can be equivalently written in the more familiar form of nested $\min - \max$ over time periods, where the dispatch decision variables are vectors rather than policy functions. See [91, Chapter 2] for more discussions. Eq. (42b) represents all the commitment constraints, including as start-up and shut-down constraints and minimum up and down times in the set X (see e.g. [79] for details on the formulation). Eq. (42c) enforces output limits when generators are on, and zero output when they are off. Eq. (42d) is the ramping constraints. Eq. (42e) restricts the dispatch level of a renewable unit to be bounded by the available power of that unit. Eqs. (42f)-(42g) are discharge and charge limit constraints for storage units. Eq. (42h) enforces energy storage capacity bounds for storage units. Eq. (42i) describes transmission line flow limit constraints. Eq. (42j) enforces system energy balance. Notice that Eqs. (42c)-(42j) are robust constraints, namely, they must hold for all $\overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r$.

In multistage robust UC model (42), the dispatch decision $\boldsymbol{p}_t(\overline{\boldsymbol{p}}^r_{[t]})$ at time t only depends on $\overline{\boldsymbol{p}}^r_{[t]}$, the realization of uncertain available renewable power *up to* time t, where $[t] := \{1, \ldots, t\}$. This makes the decision policy non-anticipative. In contrast, a *two-stage* robust UC can be formulated similarly to (42) but replacing $\boldsymbol{p}_t(\overline{\boldsymbol{p}}^r_{[t]})$ by $\boldsymbol{p}_t(\overline{\boldsymbol{p}}^r)$, thus making the dispatch decision at time t adaptive to the realization of uncertain available renewable power over all time periods, which violates non-anticipativity. A more illuminating way to see the difference is discussed in [69, Section 3.1], where the multistage robust UC is reformulated as a nested sequence of T stages of min-max problems, while the two-stage model can be reformulated as a min-max-min problem.

4.2.2 Affine Dispatch Policy

Problem (42) is quite computationally challenging, which can be seen from the fact that decision $\boldsymbol{p}(\cdot)$ is in the infinite dimensional space of functions. Alternatively, the computational difficulty can also be appreciated from the nested reformulations mentioned above.

To make the problem computationally tractable, we restrict our attention to affine policies, also known as linear decision rules [7, 61, 51]. To give some insight on the affine dispatch policy, let us write it as $p^g_{it}(\overline{\boldsymbol{p}}^r_{[t]}) = \hat{p}^g_{it} + \sum_{j \in \mathcal{N}^r} \sum_{\tau \leq t} \alpha_{itj\tau}(\overline{p}^r_{j\tau} - \hat{p}^r_{j\tau})$, where $\hat{p}^r_{j\tau}$ is the forecast (or nominal) available renewable power, $\hat{p}^g_{it}$ is the dispatch level if the realized available renewable power is equal to the forecast, and $\alpha_{itj\tau}$ is the sensitivity coefficient of dispatch on the deviation between the realized and forecast available renewable power. Note that this affine policy depends on the uncertainty at all buses and all time periods prior to t. We call such a policy a full affine policy.

It turns out that, for large-scale power systems, the above full affine policy for problem (42) is too computationally difficult to solve [69]. To deal with this difficulty, we use the following simplified affine policy that adapts to an aggregation of

uncertainty:

$$\text{Conventional units:} \quad p^g_{it}(\overline{\boldsymbol{p}}^r_{[t]}) = w^g_{it} + W^g_{it} \sum_{j \in \mathcal{N}^r} \overline{p}^r_{jt} \tag{43a}$$

$$\text{Storage units:} \quad p^{s+}_{it}(\overline{\boldsymbol{p}}^r_{[t]}) = w^{s+}_{it} + W^{s+}_{it} \sum_{j \in \mathcal{N}^r} \overline{p}^r_{jt} \tag{43b}$$

$$p^{s-}_{it}(\overline{\boldsymbol{p}}^r_{[t]}) = w^{s-}_{it} + W^{s-}_{it} \sum_{j \in \mathcal{N}^r} \overline{p}^r_{jt} \tag{43c}$$

$$\text{Renewable units:} \quad p^r_{it}(\overline{\boldsymbol{p}}^r_{[t]}) = w^r_{it} + W^r_t \, \overline{p}^r_{it}, \tag{43d}$$

where $\boldsymbol{w}, \boldsymbol{W}$ are the new decision variables. In this type of policy, dispatch decisions of generators and storage units depend linearly on the total available renewable power at a system level, and renewable units depend linearly on their own local available power. As we will show, this simplified affine policy performs surprisingly well for (42).

Problem (42) under affine policy (43) can be written in a compact form as

$$\min_{\boldsymbol{x} \in X, \boldsymbol{w}, \boldsymbol{W}} \quad \left\{ \boldsymbol{c}^\top \boldsymbol{x} \; + \; \max_{\overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r} \sum_{t \in \mathcal{T}} \boldsymbol{C}_t^\top (\boldsymbol{w}_t + \boldsymbol{W}_t \, \overline{\boldsymbol{p}}^r_t) \right\} \tag{44a}$$

$$\text{s.t.} \;\; \boldsymbol{w}_t + \boldsymbol{W}_t \, \overline{\boldsymbol{p}}^r_t \in \Omega_t \left(\boldsymbol{x}, \, \boldsymbol{w}_{[t-1]} + \boldsymbol{W}_{[t-1]} \, \overline{\boldsymbol{p}}^r_{[t-1]}, \, \overline{\boldsymbol{p}}^r_t \right) \quad \forall \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r, t \in \mathcal{T} \tag{44b}$$

where (44b) represents (42c)-(42j). Note that (44) is still a large-scale robust optimization problem with mixed-integer variables. In section 4.4, we will show that exploiting the structure of (44) is crucial to efficiently solving it.

4.2.3 Policy-guided look-ahead ED method

The solution of problem (44) not only provides a UC schedule $\boldsymbol{x}$, but also a dispatch policy $\boldsymbol{p}(\cdot)$ that may be utilized in the real-time hour-to-hour dispatch process. We will show that using this policy can significantly improve the flexibility of real-time dispatch.

Consider the real-time dispatch operation at time t. Denote the available renewable power realized up to this time as $\overline{\boldsymbol{p}}^{r,realized}_{[t]}$. For future time periods $\tau > t$, the

available renewable power is forecasted as $\overline{\boldsymbol{p}}_{\tau}^{r,forecast}$. Moreover, denote dispatch decisions of all the units realized up to time $t-1$ as $\boldsymbol{p}_{[t-1]}^{realized}$. We propose a new ED model, called the *policy-guided look-ahead* ED:

$$\min_{\hat{\boldsymbol{p}}_t,\ldots,\hat{\boldsymbol{p}}_{t+T'}} \sum_{\tau=t}^{t+T'} \sum_{i\in\mathcal{N}_g} C_i^g\, \hat{p}_{i\tau}^g \tag{45a}$$

$$\text{s.t. } \hat{\boldsymbol{p}}_t \in \Omega_t(\boldsymbol{p}_{[t-1]}^{realized}, \overline{\boldsymbol{p}}_t^{r,realized}) \tag{45b}$$

$$\hat{\boldsymbol{p}}_\tau \in \Omega_\tau(\hat{\boldsymbol{p}}_{[\tau-1]}, \overline{\boldsymbol{p}}_\tau^{r,forecast}) \;\; \tau \in [t+1:t+T'] \tag{45c}$$

$$-RD_{i,t+1}\, x_{i,t+1}^o - SD_{i,t+1}\, x_{i,t+1}^- \le p_{i,t+1}^g(\overline{\boldsymbol{p}}_{[t+1]}^r) - \hat{p}_{it}^g \le RU_{i,t+1}\, x_{it}^o + SU_{i,t+1}\, x_{i,t+1}^+ \tag{45d}$$

$$\forall\, \overline{\boldsymbol{p}}_{[t+1]}^r \in \overline{\mathcal{P}}_{[t+1]}^r \left(\overline{\boldsymbol{p}}_{[t]}^{r,realized}\right),\, i \in \mathcal{N}^g$$

$$\sum_{k=t}^{\tau} \left(\eta_i^s\, p_{ik}^{s-}(\overline{\boldsymbol{p}}_{[k]}^{r,forecast}) - p_{ik}^{s+}(\overline{\boldsymbol{p}}_{[k]}^{r,forecast})\right) = \sum_{k=t}^{\tau} \left(\eta_i^s\, \hat{p}_{ik}^{s-} - \hat{p}_{ik}^{s+}\right) \tag{45e}$$

$$\forall\, i \in \mathcal{N}^s,\, \tau \in [t+1:t+T'].$$

Here $\hat{\boldsymbol{p}}_t$ is the dispatch decision to be implemented at time t; $\hat{\boldsymbol{p}}_\tau$ for $\tau > t$ is dispatch decision for future time periods under the forecast condition; T' is the number of look-ahead time periods in the ED model; $\Omega_t(\boldsymbol{p}_{[t-1]}^{realized}, \overline{\boldsymbol{p}}_t^{r,realized})$ in (45b) represents all deterministic dispatch constraints at time t (that is, a deterministic version of eqs. (42c)-(42j) for time t); $\Omega_\tau(\hat{\boldsymbol{p}}_{[\tau-1]}, \overline{\boldsymbol{p}}_\tau^{r,forecast})$ in (45c) represents all dispatch constraints at time $\tau > t$ which depends on dispatch decisions in previous periods, $\hat{\boldsymbol{p}}_{[\tau-1]}$, and on forecast available renewable power, $\overline{\boldsymbol{p}}_\tau^{r,forecast}$.

The key constraints are (45d) and (45e). Constraint (45d) enforces *robust* ramping requirement for dispatch from time t to $t+1$, where dispatch at $t+1$ uses the affine policy $p_{i,t+1}^g(\overline{\boldsymbol{p}}_{[t+1]}^r)$ obtained from the multistage UC model (44). Constraint (45e) enforces that the energy levels of storage units under $\hat{\boldsymbol{p}}$ at time $\tau > t$ match those determined by the affine dispatch policy for forecast available renewable power. $\overline{\mathcal{P}}_{[t+1]}^r(\overline{\boldsymbol{p}}_{[t]}^{r,realized})$ is the uncertainty set $\overline{\mathcal{P}}^r$ restricted to the observed available renewable power up to time t, and projected up to time period $t+1$. The philosophy

of constraints (45d) and (45e) is that using the affine dispatch policy $p^g_{i,\tau}(\overline{\boldsymbol{p}}^r_{[\tau]})$ and $p^s_{i\tau}(\overline{\boldsymbol{p}}^{r,forecast}_{[\tau]})$ at future time periods can provide effective guidance to the dispatch decision at time t, because these affine policies are obtained from the multistage robust UC model which has a holistic view of uncertainty. We will show that this policy-guided look-ahead ED can better utilize storage devices and has a better capability to hedge against uncertainty than a deterministic look-ahead ED model (see Section 4.5.2).

4.3 Dynamic Uncertainty Set for Wind and Solar Power

Wind and solar power present significant temporal and spatial correlations [110]. In this section, we propose a new type of dynamic uncertainty sets to capture such correlations.

4.3.1 Mathematical Formulation

The dynamic uncertainty set for the available wind and solar power over a time horizon $\mathcal{T}$ is given as

$$\overline{\mathcal{P}}^r = \Big\{\overline{\boldsymbol{p}}^r = (\overline{p}^r_{it})_{i,t} \,:\, \exists\, \boldsymbol{u},\, \boldsymbol{v} \;\text{ s.t.} \tag{46a}$$

$$\overline{p}^r_{it} = f_{it} + g_{it}\, u_{it} \quad \forall\, i \in \mathcal{N}^r,\, t \in \mathcal{T} \tag{46b}$$

$$\boldsymbol{u}_t = \sum_{l=1}^{L} \boldsymbol{A}^l\, \boldsymbol{u}_{t-l} + \boldsymbol{B}\, \boldsymbol{v}_t \quad \forall\, t \in \mathcal{T} \tag{46c}$$

$$\|\boldsymbol{v}_t\| \le \Gamma \quad \forall\, t \in \mathcal{T} \tag{46d}$$

$$\sum_{t \in \mathcal{T}} \|\boldsymbol{v}_t\| \le \rho\, \Gamma\, |\mathcal{T}| \tag{46e}$$

$$0 \le \overline{p}^r_{it} \le \overline{p}^{r,max}_{it} \quad \forall\, i \in \mathcal{N}^r,\, t \in \mathcal{T} \Big\}, \tag{46f}$$

where $\overline{p}^r_{it}$ is available power of renewable unit i at time t, $\boldsymbol{f}$ and $\boldsymbol{g}$ account for deterministic seasonal components, $\|\cdot\|$ is a norm, $\Gamma > 0$ is a size parameter, $\rho \in (0, 1]$ determines a "budget over time periods", and $\overline{p}^{r,max}_{it}$ determines an upper bound on $\overline{p}^r_{it}$. Here $\boldsymbol{v}_t \in \mathbb{R}^{N_v}$ with N_v between 1 and $|\mathcal{N}^r|$.

The key feature of (46) is that both temporal and spatial correlations between uncertain wind and solar power are captured in (46c), where $\boldsymbol{A}^l$ and $\boldsymbol{B}$ determine the temporal and spatial correlations of uncertain renewable power. More specifically, $\boldsymbol{u}_t$ is the uncertainty in the wind and solar power output after the seasonality pattern (f_{it}, g_{it}) is filtered out; $\boldsymbol{u}_t$ includes both temporal and spatial correlations of these uncertain resources in $\boldsymbol{A}^l$ and $\boldsymbol{B}$ matrices; $\boldsymbol{v}_t$ represents residual uncertainty after temporal and spatial correlations are further removed from $\boldsymbol{u}_t$. So $\boldsymbol{v}_t$ can be viewed as representing a random vector with uncorrelated components over time and space. The support of $\boldsymbol{v}_t$ is described by (46d)-(46e). The size of the support is controlled by Γ, which is analogous to the maximum number of standard deviations that we allow for variations in each component of $\boldsymbol{v}_t$. We call (46) a *dynamic* uncertainty set. As a special case, if the dimension of $\boldsymbol{v}_t$ is the same as that of $\boldsymbol{u}_t$, $\boldsymbol{B}$ is the identity matrix, $\boldsymbol{A}^l$'s are zero, and $\rho = 1$, we obtain a *static* uncertainty set that ignores temporal and spatial correlations and is separable over time periods, similar to the budget uncertainty set used in literature (see e.g. [13, 69]).

To further illustrate the meaning of Γ, consider the static case described above (where $\boldsymbol{u}_t = \boldsymbol{v}_t$ and $\rho = 1$), the norm is the ℓ_∞ norm, i.e. $\|\cdot\| = \|\cdot\|_\infty$, and g_{it} corresponds to the standard deviation of available power at renewable unit i and time t. In this case we obtain $f_{it} - \Gamma\, g_{it} \leq \overline{p}^r_{it} \leq f_{it} + \Gamma\, g_{it}$ and no coupling relations between renewable units or time periods, and we can thus interpret Γ as the "number of standard deviations" that we allow for available power variations at each renewable unit and time period. This is the most basic uncertainty set corresponding to a box. For the general dynamic uncertainty set, the intuition for the choice of Γ is similar, except that Γ will now determine the "number of standard deviations" that we allow for variations of each component in $\boldsymbol{v}_t$.

The concept of dynamic uncertainty sets was first proposed in [67], where uncertainty in *wind speed* is modeled and a power curve is used to transform wind speed into

available wind power. In this Chapter, we directly model uncertainty in renewable power to improve computational efficiency. Furthermore, we model both wind and solar power uncertainty and their correlations. Notice that the norm in (46d), (46e) can be any norm such as $\ell_1, \ell_2, \ell_\infty$ or a combination thereof, such as the intersection of ℓ_1 and ℓ_∞.

4.3.2 Parameter Estimation And Dimension Reduction

To estimate the parameters of the dynamic uncertainty set (46), we consider the following stochastic model for available power of renewable units:

$$\tilde{p}^r_{it} = f_{it} + g_{it}\,\tilde{u}_{it} \quad \forall\, i \in \mathcal{N}^r,\, t \in \mathcal{T} \tag{47a}$$

$$\tilde{\boldsymbol{u}}_t = \sum_{l=1}^{L} \boldsymbol{A}^l\,\tilde{\boldsymbol{u}}_{t-l} + \tilde{\boldsymbol{\epsilon}}_t \quad \forall\, t \in \mathcal{T}, \tag{47b}$$

where $\tilde{p}^r_{it}$ is the power available at renewable unit i at time t, and $\tilde{\boldsymbol{\epsilon}}_t$ is an i.i.d. random vector under certain distribution. In this Chapter, we assume $\tilde{\boldsymbol{\epsilon}}_t$ to have a multivariate normal distribution centered at zero with covariance matrix $\boldsymbol{\Sigma}$. Other distributions could be assumed with a proper choice of the norm in (46d), (46e) in the dynamic uncertainty set (e.g. using the approach in [22]).

In order to estimate the parameters of this model, $\boldsymbol{f}$ and $\boldsymbol{g}$ can be estimated using linear regression after identifying daily seasonality in the volatility of available renewable power. Next, $\tilde{\boldsymbol{u}}_t$ corresponds to a multivariate autoregressive process, and given a choice of time lag L, $\boldsymbol{A}^l$ and $\boldsymbol{\Sigma}$ can be estimated using statistical inference techniques from time series analysis [86]. Using this, $\boldsymbol{B}$ in (46) can be estimated by Cholesky decomposition of $\boldsymbol{\Sigma}$ as $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}^\top$.

At this point, it is important to note that the dimension of the uncertainty set plays a fundamental role in the difficultness of solving the associated robust optimization problem. Under the presence of many renewable units, $\overline{\mathcal{P}}^r$ can have a large dimension. We can reduce the dimension of $\overline{\mathcal{P}}^r$ by principal component analysis as follows. The matrix $\boldsymbol{\Sigma}$ can be eigen-decomposed as $\boldsymbol{\Sigma} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^\top$, where $\boldsymbol{V}$ contains

the eigenvectors and $\boldsymbol{\Lambda}$ has the eigenvalues in the diagonal. Then, we ignore the smaller eigenvalues in $\boldsymbol{\Lambda}$ and the corresponding eigenvectors in $\boldsymbol{V}$ by removing the corresponding columns of $\boldsymbol{B} = \boldsymbol{V}\boldsymbol{\Lambda}^{1/2}$. In this way, the number of columns N_v in (46) left in $\boldsymbol{B}$ can be any number from 1 to $|\mathcal{N}^r|$. If N_v is selected too close to $|\mathcal{N}^r|$, then a high-dimensional uncertainty set is obtained, resulting in a large problem. If N_v is too close to 1, then the uncertainty representation may be too inaccurate. The right balance will depend on the particular instance solved.

4.4 Solution Method

The affine multistage robust UC model (44) is a so-called semi-infinite program, i.e. there are finite number of decision variables, but infinite number of constraints. Due to this, a deterministic counterpart of (44) needs to be formulated. There are two main classes of approaches for this purpose. The most widely used approach is the duality based approach [6] that replaces each robust constraint by its dual program with additional variables and constraints. The other less explored approach is based on constraint generation [10, 69], which dynamically generates violated scenarios and the associated deterministic constraints. These two approaches are combined in the solution method proposed in this section, and special structures of (44) are exploited.

Section 4.4.1 introduces the basic constraint generation framework. Section 4.4.2 exploits the special structure of the robust generation limit and energy balance constraints in (44). Section 4.4.3 presents an outer approximation method for reformulating the inter-temporal constraints in (44). Section 4.4.4 discusses several techniques to further enhance the efficiency of the overall algorithm.

4.4.1 Constraint Generation Framework

Constraint generation (CG) is recently applied to solve large-scale robust optimization problems [69, 10]. The master robust UC problem in the CG algorithm can be written

as

$$\min_{(\boldsymbol{x},\boldsymbol{w},\boldsymbol{W},z)\in\Omega} \quad \boldsymbol{c}^\top \boldsymbol{x} + z \tag{48a}$$

$$\text{s.t. } \boldsymbol{a}_k(\boldsymbol{W})^\top \overline{\boldsymbol{p}}^r \le b_k(\boldsymbol{x},\boldsymbol{w},z) \ \forall\, k \in [K],\, \overline{\boldsymbol{p}}^r \in P_k, \tag{48b}$$

where z corresponds to the worst-case dispatch cost, K is the number of robust constraints in the problem, and P_k is the current set of extreme points of the uncertainty set $\overline{\mathcal{P}}^r$ considered for the k-th robust constraint. The basic CG algorithm solves the master problem (70), and checks if the k-th robust constraint is violated by the current solution, and if so, the associated worst-case scenario from $\overline{\mathcal{P}}^r$ is added to P_k, and the master program is solved again. This method is formally presented in Algorithm 5.

Algorithm 5 Constraint generation

1: **repeat**
2: Solve Master Problem (70)
3: **for all** $k \in \{1, ..., K\}$ **do**
4: $\overline{\boldsymbol{p}}^r_k \leftarrow \text{argmax}\left\{\boldsymbol{a}_k(\boldsymbol{W})^\top \overline{\boldsymbol{p}}^r : \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r\right\}$
5: If $\boldsymbol{a}_k(\boldsymbol{W})^\top \overline{\boldsymbol{p}}^r_k > b_k(\boldsymbol{x},\boldsymbol{w},z)$ let $P_k \leftarrow P_k \cup \{\overline{\boldsymbol{p}}^r_k\}$
6: **end for**
7: **until** $\boldsymbol{a}_k(\boldsymbol{W})^\top \overline{\boldsymbol{p}}^r_k \le b_k(\boldsymbol{x},\boldsymbol{w},z)$ for all $k \in [K]$

To give a concrete example of the robust constraint (48b), let us consider the upper output limit constraint in (42c) for a generator i at time t under affine policy (43). The robust constraint in this case is given by

$$w^g_{it} + W^g_{it} \sum_{j\in\mathcal{N}^r} \overline{p}^r_{jt} \le x^o_{it}\, \overline{p}^g_{it} \quad \forall \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r,$$

where $a_{kj\tau}(\boldsymbol{W}) = W^g_{it}$ for $\tau = t$ and $a_{kj\tau}(\boldsymbol{W}) = 0$ for $\tau \neq t$, for any $j \in \mathcal{N}^r$, and $b_k(\boldsymbol{x},\boldsymbol{w},z) = x^o_{it}\, \overline{p}^g_{it} - w^g_{it}$. All other robust constraints in the master problem are similarly defined.

This basic CG framework is the starting point to develop a practical algorithm for solving large-scale robust UC problems. The key is to fully exploit the structure

of (44). In the full algorithm described in this section, we will handle transmission line flow limit constraints (42i) through CG; for all other constraints, we use more efficient reformulations.

4.4.2 Reformulation of generation limit and balance constraints

The deterministic counterparts of the robust generation limit (42c) and the energy balance constraints (42j) in the affine UC model (44) can be explicitly derived without any dualization of the uncertainty set.

4.4.2.1 Generation limit constraints

Due to the structure of the simplified affine policy (43), we can directly identify the worst-case scenarios for the robust generation limit constraints.

Proposition 9. *Under affine policy* (43)*, robust generation limit constraints* (42c) *are equivalent to*

$$x^o_{it}\,\underline{p}^g_{it} \le w^g_{it} + W^g_{it}\,\overline{p}^{r,total}_t \le x^o_{it}\,\overline{p}^g_{it} \quad \forall\, i \in \mathcal{N}^g,\, t \in \mathcal{T},\, \overline{p}^{r,total}_t \in \left\{\overline{p}^{r,total,min}_t,\, \overline{p}^{r,total,max}_t\right\}, \tag{49}$$

where

$$\overline{p}^{r,total,min}_t = \min_{\overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r} \sum_{j \in \mathcal{N}^r} \overline{p}^r_{jt} \quad \forall\, t \in \mathcal{T} \tag{50a}$$

$$\overline{p}^{r,total,max}_t = \max_{\overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r} \sum_{j \in \mathcal{N}^r} \overline{p}^r_{jt} \quad \forall\, t \in \mathcal{T}. \tag{50b}$$

Proof of Proposition 9. The result follows from the fact that

$$\max_{\overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r} a \sum_{j \in \mathcal{N}^r} \overline{p}^r_{jt} = \max\left\{a\,\overline{p}^{r,total,min}_t,\, a\,\overline{p}^{r,total,max}_t\right\},$$

for any given a. □

A similar result holds for robust constraints of storage units' output and input limits (42f) and (42g), as well as for renewable unit output limit constraints (42e). Details are omitted for space.

4.4.2.2 Energy balance constraints

The deterministic counterpart of the robust energy balance constraints can also be obtained in closed form.

Proposition 10. *Under affine policy* (43)*, robust energy balance constraints* (42j) *are equivalent to the following system of equations, for every* $t \in \mathcal{T}$*,*

$$\sum_{i\in\mathcal{N}^g} w_{it}^g + \sum_{i\in\mathcal{N}^r} w_{it}^r + \sum_{i\in\mathcal{N}^s} \left(w_{it}^{s+} - w_{it}^{s-}\right) = \sum_{j\in\mathcal{N}^d} d_{jt} \tag{51a}$$

$$\sum_{i\in\mathcal{N}^g} W_{it}^g + W_t^r + \sum_{i\in\mathcal{N}^s} (W_{it}^{s+} - W_{it}^{s-}) = 0, \tag{51b}$$

whenever the uncertainty set $\overline{\mathcal{P}}^r$ *is full-dimensional.*

Proof of Proposition 10. The robust energy balance constraint for each t can be written compactly as $\boldsymbol{a}_t(\boldsymbol{W})^\top \overline{\boldsymbol{p}}_t^r = b_t(\boldsymbol{w})$ for all $\overline{\boldsymbol{p}}_t^r \in \overline{\mathcal{P}}^r$, where $\boldsymbol{a}_t(\boldsymbol{W})$ and $b_t(\boldsymbol{w})$ are linear in $\boldsymbol{W}$ and $\boldsymbol{w}$, respectively. If $\overline{\mathcal{P}}^r$ is full-dimensional, then $\boldsymbol{a}_t(\boldsymbol{W}) = \boldsymbol{0}$ and $b_t(\boldsymbol{w}) = 0$ must hold, which gives (51a)-(51b). □

4.4.3 Outer approximation for inter-temporal constraints

The worst-case cost constraint, ramping constraints (42d), and the storage capacity constraints (42h) all involve decisions over consecutive time periods, i.e. they induce inter-temporal coupling between dispatch decisions. Dualizing these robust constraints introduces a large number of new variables and constraints, while directly applying CG may lead to slow convergence. In this section, we introduce an outer approximation method for efficient reformulation. Observe that, due to the simplified affine policy structure (43), the inter-temporal robust constraints in (44) only depend on total system-level available renewable power rather than on bus-level details. So we can project the bus-level uncertainty set (46) to its equivalence for system-level uncertainty. However, this latter uncertainty set still involves a large number of variables. Thus, we use outer approximation (OA) to further reduce its dimension. This

technique is general. We use it here to reformulate inter-temporal constraints, and in section 4.4.4.2 for screening transmission constraints.

The inter-temporal robust constraints in (44) can be written as

$$\max_{\overline{\boldsymbol{p}}^{r,total}_{[t_1:t_2]} \in \overline{\mathcal{P}}^{r,total}_{[t_1:t_2]}} \sum_{t=t_1}^{t_2} a_t^{total}(\boldsymbol{W})\, \overline{p}_t^{r,total} \;\leq\; b(\boldsymbol{x}, \boldsymbol{w}, z), \tag{52}$$

where

$$\overline{\mathcal{P}}^{r,total}_{[t_1:t_2]} = \Big\{ \overline{\boldsymbol{p}}^{r,total}_{[t_1:t_2]} :\; \exists\, \overline{\boldsymbol{p}}^{r} \in \overline{\mathcal{P}}^{r} \text{ s.t. } \overline{p}_t^{r,total} = \sum_{j \in \mathcal{N}^r} \overline{p}^r_{jt} \;\; \forall t \in [t_1 : t_2] \Big\}$$

is the projection of bus-level uncertainty set $\overline{\mathcal{P}}^{r}$ unto the total available renewable power, and $a_t^{total}(\boldsymbol{W}), b(\boldsymbol{x}, \boldsymbol{w}, z)$ are properly defined depending on the particular robust constraint. We then replace $\overline{\mathcal{P}}^{r,total}_{[t_1:t_2]}$ by the following OA $\widehat{\mathcal{P}}^{r,total}_{[t_1:t_2]}$.

$$\begin{aligned} \widehat{\mathcal{P}}^{r,total}_{[t_1:t_2]} = \Big\{\, \overline{\boldsymbol{p}}^{r,total}_{[t_1:t_2]} :\;\; & \overline{p}_t^{r,total,min} \leq \overline{p}_t^{r,total} \leq \overline{p}_t^{r,total,max} \;\; \forall t \in [t_1 : t_2] \\ & \underline{\triangle}_t^{total} \leq \overline{p}_t^{r,total} - \overline{p}_{t-1}^{r,total} \leq \overline{\triangle}_t^{total} \;\; \forall t \in [t_1 + 1 : t_2] \,\Big\}, \end{aligned} \tag{53}$$

where

$$\underline{\triangle}_t^{total} = \min_{\overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r} \sum_{j \in \mathcal{N}^r} \left(\overline{p}^r_{jt} - \overline{p}^r_{j,t-1} \right) \;\forall\, t \in [t_1 + 1 : t_2] \tag{54a}$$

$$\overline{\triangle}_t^{total} = \max_{\overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r} \sum_{j \in \mathcal{N}^r} \left(\overline{p}^r_{jt} - \overline{p}^r_{j,t-1} \right) \;\forall\, t \in [t_1 + 1 : t_2]. \tag{54b}$$

Observe that this set is indeed an OA of $\overline{\mathcal{P}}^{r,total}_{[t_1:t_2]}$ (i.e. $\overline{\mathcal{P}}^{r,total}_{[t_1:t_2]} \subset \widehat{\mathcal{P}}^{r,total}_{[t_1:t_2]}$) due to (54), thus ensuring robust feasibility when replacing $\overline{\mathcal{P}}^{r,total}_{[t_1:t_2]}$ in (52). Also observe that the OA (53) only involves the $\overline{p}_t^{r,total}$ variables, whereas $\overline{\mathcal{P}}^{r,total}_{[t_1:t_2]}$ has many more additional variables $\boldsymbol{u}, \boldsymbol{v}$ as in (46). Thus, solving (52) over the OA set becomes much faster than over the original uncertainty set $\overline{\mathcal{P}}^{r,total}_{[t_1:t_2]}$. The proposition below summarizes this result, and the use of duality based approach to reformulate the resulting robust constraint.

Proposition 11. *Robust constraint* (52) *is implied by*

$$\max_{\overline{\boldsymbol{p}}^{r,total}_{[t_1:t_2]} \in \widehat{\mathcal{P}}^{r,total}_{[t_1:t_2]}} \sum_{t=t_1}^{t_2} a_t^{total}(\boldsymbol{W})\, \overline{p}_t^{r,total} \;\leq\; b(\boldsymbol{x}, \boldsymbol{w}, z),$$

which is equivalent to the existence of vectors $\overline{\boldsymbol{\pi}}, \underline{\boldsymbol{\pi}}, \overline{\boldsymbol{\phi}}, \underline{\boldsymbol{\phi}} \geq 0$ *such that*

$$\sum_{t=t_1}^{t_2} \left(\overline{p}_t^{r,total,max}\, \overline{\pi}_t - \overline{p}_t^{r,total,min}\, \underline{\pi}_t \right) + \sum_{t=t_1+1}^{t_2} \left(\overline{\triangle}_t^{total}\, \overline{\phi}_t - \underline{\triangle}_t^{total}\, \underline{\phi}_t \right) \;\leq\; b(\boldsymbol{x}, \boldsymbol{w}, z)$$

$$\overline{\pi}_t - \underline{\pi}_t + \overline{\phi}_t - \underline{\phi}_t - \overline{\phi}_{t+1} + \underline{\phi}_{t+1} = a_t^{total}(\boldsymbol{W}) \quad \forall t \in [t_1 : t_2]$$

$$\overline{\phi}_{t_1} = \underline{\phi}_{t_1} = \overline{\phi}_{t_2+1} + \underline{\phi}_{t_2+1} = 0.$$

Proof of Proposition 11. The result follows from $\overline{\mathcal{P}}^{r,total}_{[t_1:t_2]} \subset \widehat{\mathcal{P}}^{r,total}_{[t_1:t_2]}$ and then taking the dual over $\widehat{\mathcal{P}}^{r,total}_{[t_1:t_2]}$. □

Finally, we note that in [69] the separability over time periods of a static uncertainty set is exploited to reformulate ramping constraints in a simple way, also exploiting the fact that these constraints only couple two time periods. However, the more general dynamic uncertainty sets (46) are not separable over time periods, and energy storage constraints couple up to T time periods. Due to this, the technique presented above based on OA and the duality based approach is a critical enhancement that allows efficiently handling energy storage and dynamic uncertainty sets in the multistage robust UC.

See [99, Chapter 6] for more discussion on outer approximations.

4.4.4 Further Algorithmic Enhancements

The CG framework is further enhanced by the following techniques.

4.4.4.1 One-tree Benders implementation

The presence of binary variables in the master problem implies that if the constraint generation approach presented in 4.4.1 is directly used, then a potentially large number of mixed-integer programs will have to be solved throughout the algorithm, which

may be quite slow. We propose an efficient alternative which is to use the one-tree Benders approach, in which the solver builds only one branch-and-bound tree and adds the generated constraints as the branch-and-bound process unfolds. Solver callbacks are required in this implementation. For example, lazy constraint callbacks [48] are required in CPLEX.

Another alternative, discussed in [69], consists of "fixing and releasing" binary variables in order to generate multiple cuts to the mixed-integer master problem. However, this still requires potentially building from scratch more than one branch-and-bound tree and is dominated by the more efficient one-tree Benders approach.

4.4.4.2 Constraint screening using fast computed upper bounds

Each iteration of the CG algorithm needs to solve the following separation problem for each robust constraint k, by checking whether

$$\max_{\overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r} \ \boldsymbol{a}_k(\boldsymbol{W})^\top \overline{\boldsymbol{p}}^r \ \leq \ b_k(\boldsymbol{x}, \boldsymbol{w}, z) \tag{55}$$

holds for the fixed $z, \boldsymbol{x}, \boldsymbol{w}, \boldsymbol{W}$. It amounts to solving a linear program over the dynamic uncertainty set (46) for each k. This can be time consuming for large instances. We propose to screen the robust constraints in the following way. In each iteration of the CG master problem, calculate an *upper bound* $ub_k(\boldsymbol{W})$ for the left-hand side of (55) as

$$ub_k(\boldsymbol{W}) = \max_{\overline{\boldsymbol{p}}^r \in \widehat{\mathcal{P}}_k^r} \ \boldsymbol{a}_k(\boldsymbol{W})^\top \overline{\boldsymbol{p}}^r, \tag{56}$$

where $\widehat{\mathcal{P}}_k^r$ is an outer approximation of $\overline{\mathcal{P}}^r$, so that $\max_{\overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r} \ \boldsymbol{a}_k(\boldsymbol{W})^\top \overline{\boldsymbol{p}}^r \ \leq \ ub_k(\boldsymbol{W})$. Then, before solving (55) for each robust constraint k, we check whether $ub_k(\boldsymbol{W}) \leq b_k(\boldsymbol{w}, z)$. If this holds, then (55) must also hold, and we do not need to solve a linear program over the uncertainty set. Otherwise, we solve the linear program to check if (55) holds or not.

In order for this screening process to be efficient, (56) must be solved very fast. For this purpose, we construct interval type sets $\widehat{\mathcal{P}}_k^r$ for robust transmission constraints, under which upper bounds can be computed by simply checking the sign of the elements in $\boldsymbol{W}$. Since typically many robust constraints are far from being violated (e.g., some transmission lines are rarely congested), such robust constraints will be screened very rapidly using this technique.

4.4.4.3 Strategy for checking loose constraints

After checking for the feasibility of all robust constraints in the problem, it is possible that several of them are quite far from being violated. These loose constraints are unlikely to become violated in the next iteration. Given this, we restrict the set of robust constraints to those that were violated or close to being violated in the last iteration, and only get back to checking all robust constraints once the master problem has converged. This process can be repeated as needed until the optimal solution has been found and feasibility is ensured for all robust constraints.

4.5 Computational Experiments

We conduct extensive computational experiments to evaluate the solution method and to understand the benefits of the new model. The experiments are carried out using an adapted version of the 2736-bus Polish system [120]. The system contains 289 generators (28880 MW of total capacity), 60 wind farms (10689 MW installed), 30 solar farms (6299 MW installed), 10 storage units (600 MW of total output capacity), 2011 demand nodes (17831 MW average, 22594 MW peak) and 100 transmission lines. The energy storage capacity of each storage unit corresponds to five hours at full output, their ramping capacities are unconstrained, and their efficiency is 80% [30]. We use wind and solar power data from NREL's Western Wind and Solar Integration Datasets [85]. All the experiments have been implemented using Python 2.7 in a PC with an Intel Core i5 processor at 2.4 GHz with 4GB memory, using

CPLEX 12.6 as the MIP solver. Section 4.5.1 studies the performance of the proposed solution method, including a comparison of solution quality obtained with and without the outer-approximation technique. Section 4.5.2 evaluates the advantages of the proposed approach.

4.5.1 Performance of the solution method

Here we study how the various techniques presented in section 4.4 contribute to an effective solution method. In the experiments, we use a horizon of $T = 24$ hours with hourly intervals. The $f, g, \boldsymbol{A}, \boldsymbol{B}$ parameters of the dynamic uncertainty set (46) are estimated using 30 days of the NREL data. The time lag is set as $L = 1$, and the dimension of $\boldsymbol{v}$ is set as $N_v = 25$, which offers a good balance between uncertainty representability and computational tractability. For the size parameters, we set $\rho = 0.1$ and test various Γ values. The norm in (46c)-(46d) is defined as $\|\cdot\|_{l_1 \cap l_\infty} = \max\left\{\|\cdot\|_1/\sqrt{N_v}\,,\ \ \|\cdot\|_\infty\right\}$, resulting in a polyhedral uncertainty set. An optimality gap of 1% is used for solving MIPs in the CG algorithm.

Table 16: Solution time (hours) for the 2736-bus system

Γ	0.25	0.5	1	2	3	4
CG	T	T	T	T	T	T
CG + OTB	T	T	T	T	T	T
CG + OTB + OA	1.53	2.16	1.50	1.79	1.59	2.64
CG + OTB + OA + CS	0.96	2.04	1.27	1.45	1.30	0.79

Table 16 presents the computation time of different combinations of solution techniques presented in Section 4.4. In this Table, "CG" corresponds to the basic constraint generation algorithm described in Section 4.4.1, "OTB" is the one-tree Benders implementation discussed in Section 4.4.4.1, "OA" uses the outer-approximation technique in Section 4.4.3, and "CS" uses constraint screening and the strategy for checking loose constraints in Sections 4.4.4.2 and 4.4.4.3. All the methods incorporate the reformulations in Section 4.4.2, so "CG+OTB+OA+SC" is the full solution method presented in Section 4.4. "T" stands for reaching a time limit of 6 hours.

Table 17: Worst-case cost for the 2736-bus system with and without outer approximations

Γ	0.25	0.5	1	2	3	4
With OA (M$)	11.675	11.892	12.368	13.194	13.833	14.428
Without OA (M$)	11.675	11.881	12.344	13.115	13.738	14.324
Difference	0.00%	0.09%	0.19%	0.61%	0.69%	0.72%

We can observe that all of the enhancements are important for efficiency. The basic CG has to solve many difficult mixed integer programs, making it very slow for such a large-scale instance. Method "OTB" builds only one branch-and-bound tree and generates the constraints in an integrated way, however, the number of robust constraints generated through CG is still very large, making the process slow. This is fixed in "CG+OTB+OA" by reformulating the robust inter-temporal constraints for worst-case cost, ramping, and storage capacities, in such a way that they are enforced throughout the whole progress of the algorithm with a simple computational representation, leaving the sequential generation of constraints only for transmission. Finally, "CG+OTB+OA+SC" further improves the algorithm by reducing the overall number of separation problems (55) solved, through quickly recognizing several robust constraints that are not violated.

To further show the effectiveness of outer approximation, we study its tightness. In particular, Table 17 compares the worst-case cost obtained using the outer approximation method in section 4.4.3 to that obtained without using this technique (with running time longer than 6 hours). We can observe that the loss of solution quality is small, specially for small values of the size parameter Γ, confirming the value of this technique.

We also study whether the affine policy proposed (43) is a close-to-optimal solution for the fully-adaptive multistage robust UC model (42). For this purpose, we calculate the worst-case cost obtained by a corresponding two-stage robust UC

Table 18: Worst-case cost under two-stage robust UC and gap obtained for multistage robust UC with and without outer approximations, for the 2736-bus system

Gamma	0.25	0.5	1	2	3	4
Two-Stage (M$)	11.578	11.777	12.193	12.850	13.443	13.884
Gap With OA	0.84%	0.98%	1.44%	2.68%	3.13%	3.92%
Gap Without OA	0.84%	0.88%	1.24%	2.06%	2.19%	3.17%

model with a min-max-min structure, in order to obtain a lower bound to the optimal worst-case cost of the fully-adaptive multistage robust UC. The two-stage robust UC is a relaxation of the fully-adaptive multistage robust UC because it relaxes the non-anticipativity of dispatch decisions [69]. With this, we calculate the gap between the worst-case cost obtained under the affine policy, with and without outer approximations, and the worst-case cost obtained under the two-stage robust UC, which provides a sense of whether the affine policy provides a good solution or not for the fully-adaptive multistage robust UC. Table 18 shows the corresponding results. We can observe that under small values of the size parameter Γ, the affine policy obtains good solutions with a gap smaller than 1%, loosing some quality when this parameter is larger.

Given the complexity of the multistage robust UC with dynamic uncertainty sets, the large-scale 2736-bus instance solved here, and the simple computer where these experiments were carried out, we believe that the solution method proposed here is very promising for an eventual practical implementation in real-world power systems with a significant adoption of wind and solar power.

4.5.2 Comparison to other UC and ED models

This section studies the performance of three different UC solutions and ED methods on a simulation platform of the dispatch process that mimics the hour to hour operation of the power system. This simulation consists of a rolling-horizon process where, given an on/off schedule for generators (UC solution $\boldsymbol{x}$), a dispatch problem is solved

for every $t = 1, \ldots, T$, starting with $t = 1$ and moving forward until $t = T$, with uncertain parameters at time t revealed only at that time. That is, when solving a dispatch problem at time t the values of uncertain parameters at future time periods are not known. The dispatch problem solved at time t implements dispatch decisions for that time, and it takes as input the dispatch decisions implemented in the previous time periods. $N = 100$ such simulations are carried out, with $T = 24$ hours, and then several cost and reliability metrics are examined. The trajectories for wind and solar power are generated using the stochastic model in Eq. (47), using 30 days of data for parameter estimation. For the 2736-bus system, the $N = 100$ simulated trajectories present an average of 5164 MW for available wind power and 1133 MW for available solar power, resulting on an average renewable penetration of 35.3%.

The following UC and ED models are tested: multistage robust UC with dynamic uncertainty sets using the policy-guided look-ahead ED method proposed in section 4.2.3 (RobUC-Dynamic), multistage robust UC with static uncertainty sets using the policy-enforcement ED method proposed in [69] (RobUC-Static), and deterministic UC with reserves using deterministic look-ahead ED (DetUC).

The deterministic UC corresponds to a modification of problem (42) in the case where the uncertainty set only contains the forecast trajectory for available renewable power, $\overline{\mathcal{P}}^r = \{\overline{\boldsymbol{p}}^{r,forecast}\}$, thus collapsing the dispatch policy to one dispatch plan rather than a function, $\boldsymbol{p}_t(\overline{\boldsymbol{p}}^r_{[t]}) = \boldsymbol{p}_t$. This model is further enhanced by reserves through replacing eq. (42c) by

$$x^o_{it}\,\underline{p}^g_{it} + r^-_{it} \leq p^g_{it} \leq x^o_{it}\,\overline{p}^g_{it} - r^+_{it} \quad \forall i \in \mathcal{N}^g, t \in \mathcal{T},$$

and adding constraints

$$\sum_{i \in \mathcal{N}^g} r^-_{it} \geq R^-_t\,, \quad \sum_{i \in \mathcal{N}^g} r^+_{it} \geq R^+_t \quad \forall t \in \mathcal{T},$$

where $r^-_{it}, r^+_{it} \geq 0$ are the down-reserve and up-reserve provided by generator i at time t, and R^-_t, R^+_t are the down-reserve and up-reserve requirement levels at time

t. Notice that the multistage robust UC (42) does not need to consider reserve requirements, given that it addresses uncertainty in a direct and systematic way, determining reserves endogenously.

The policy-enforcement ED can be formulated as

$$\min_{\hat{\boldsymbol{p}}_t} \left\{ \sum_{i \in \mathcal{N}_g} C_i^g \, \hat{p}_{it}^g : \text{ Eqs. (45b) and (45d) hold} \right\},$$

and the deterministic look-ahead ED as

$$\min_{\hat{\boldsymbol{p}}_t, \cdots, \hat{\boldsymbol{p}}_{t+T'}} \left\{ \sum_{\tau=t}^{t+T'} \sum_{i \in \mathcal{N}_g} C_i^g \, \hat{p}_{i\tau}^g : \text{ Eqs. (45b) and (45c) hold} \right\}.$$

With this we can see that the policy-guided look-ahead ED (45) generalizes both the above EDs by utilizing the affine policies obtained from the multistage robust UC in the robust ramping constraints (45d) and enforcing storage levels constraints (45e) in a multi-period look-ahead framework, which are both very important to leverage the benefits of energy storage resources. Here, we use $T' = 3$ look-ahead periods.

For the two robust UC models, the size of the uncertainty sets is parameterized by Γ, and for the deterministic UC, the reserve requirement levels R_t^-, R_t^+ are selected as $R_t^- = R_t^+ = \Gamma \, \sigma_t^{TNL}$, where σ_t^{TNL} is the standard deviation of total net load (namely, total demand minus total available renewable power) at time t, under the simulated trajectories [17]. To properly study the performance of all these methods, ED problems are extended with penalty variables for violations of energy balance and transmission line capacity limits, each with a unit cost of \$5000/MWh.

The simulation results for the 2736-bus system are presented in Table 19, where "Cost Avg" is the average of total cost over the $N = 100$ simulations, "Cost Std" is the standard deviation of total cost, "Cost CVaR" is the conditional value at risk of total cost at a 10% level (that is, the average total cost of the 10 highest total costs, given $N = 100$), "Penalty Cost Avg" is the average penalty cost, "Penalty Freq" is the proportion of time periods where penalty occurred, "Renewables Util",

utilization of renewables, is the proportion of used renewable power with respect to available renewable power, and "Stored Avg" is the average level of stored energy.

4.5.2.1 Robust UC v.s. Deterministic UC

First, for both robust and deterministic UC models, the trade-off between operational cost and system reliability (cost std, penalty freq, CVaR) is controlled by the uncertainty set size parameter Γ and the reserve level parameter, respectively. Higher Γ or reserve level improves system reliability but may increase cost. Second, the robust UC models significantly improve both the operational cost and reliability over DetUC. In particular, comparing to the best economic performance of DetUC ($\Gamma = 4$), the RobUC-Dynamic model at $\Gamma = 1$ achieves a decrease of 7.62% in average cost, 91.64% in cost std, 35.49% in CVaR, and at the same time completely eliminates penalty; the RobUC-Static model at $\Gamma = 3$ achieves a decrease of 7.04% in average cost, 90.77% in cost std, and 34.87% in CVaR, and also eliminates penalty. Notice that, at its best performance ($\Gamma = 4$), DetUC still has substantial penalty cost. See Figure 11 for a graphical representation of this comparison. Third, the robust models curtail renewables slightly more than the DetUC model to achieve significantly improved system reliability.

4.5.2.2 RobUC-Dynamic v.s. RobUC-Static

Robust UC with dynamic uncertainty sets (RobUC-Dynamic) further improves over robust UC with static uncertainty sets (RobUC-Static). First, RobUC-Dynamic at $\Gamma = 1$ achieves better performance in all three categories: 0.62% lower cost avg, 9.57% lower cost std, and 0.96% lower CVaR than RobUC-Static at its best performance $\Gamma = 3$. In other words, robust UC with dynamic uncertainty sets *dominates* the performance of static uncertainty sets. Second, we can also observe that the average level of storage utilization for RobUC-Dynamic is much higher than that of RobUC-Static. This difference is mainly driven by the different ED models used. In particular,

Table 19: Simulation results for Polish 2736-bus system

Multistage robust UC with dynamic uncertainty set using policy-guided look-ahead ED (RobUC-Dynamic)						
Γ	0.25	0.5	1	2	3	4
Cost Avg (M$)	12.089	11.459	11.567	11.729	11.865	12.017
Cost Std (M$)	1.991	0.262	0.189	0.199	0.202	0.200
Cost CVaR (M$)	17.343	12.000	11.907	12.086	12.228	12.377
Penalty Cost Avg ($)	29424	884	0	0	0	0
Penalty Freq	4.67%	0.29%	0.00%	0.00%	0.00%	0.00%
Renewables Util	99.2%	99.1%	99.1%	98.6%	97.6%	96.5%
Stored Avg (MWh)	613	689	726	709	847	1080

Multistage robust UC with static uncertainty set using policy-enforcement ED (RobUC-Static)						
Γ	0.25	0.5	1	2	3	4
Cost Avg (M$)	24.844	18.215	13.676	11.671	11.639	11.765
Cost Std (M$)	11.654	7.998	3.851	0.608	0.209	0.213
Cost CVaR (M$)	49.390	36.601	23.397	13.037	12.023	12.156
Penalty Cost Avg ($)	563688	285140	93386	4696	0	0
Penalty Freq	31.63%	19.83%	8.63%	0.58%	0.00%	0.00%
Renewables Util	98.6%	98.6%	98.3%	97.5%	96.7%	96.1%
Stored Avg (MWh)	26	46	88	101	129	151

Deterministic UC with reserve using deterministic look-ahead ED (DetUC)						
Γ	0.25	0.5	1	2	3	4
Cost Avg (M$)	18.405	20.878	16.687	16.485	17.018	12.520
Cost Std (M$)	6.935	7.446	5.605	5.363	5.805	2.260
Cost CVaR (M$)	34.362	36.326	30.015	29.058	29.739	18.458
Penalty Cost Avg ($)	295995	400152	223927	215491	237765	48136
Penalty Freq	37.13%	32.83%	32.33%	32.08%	27.92%	9.63%
Renewables Util	98.8%	99.0%	98.8%	98.8%	98.9%	99.0%
Stored Avg (MWh)	384	380	381	381	369	377

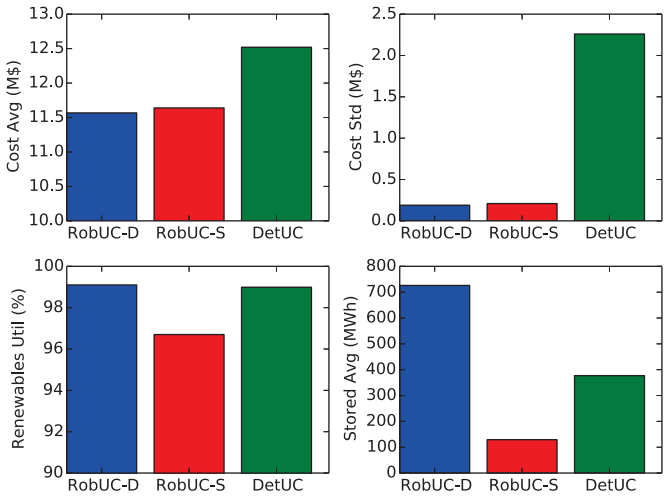

Figure 11: Performance measures for RobUC-Dynamic at $\Gamma = 1$, RobUC-Static at $\Gamma = 3$ and DetUC at $\Gamma = 4$

the policy guided look-ahead ED in RobUC-Dynamic is more effective at deploying energy storage devices than the policy-enforcement ED by following the multistage policy for storage decisions. Third, RobUC-Dynamic utilizes more renewable power than RobUC-Static for all levels of Γ. This is due to the fact that the dynamic uncertainty set is more realistic and less conservative than the static uncertainty set.

In summary, robust UC models dominate the deterministic UC model in all operational cost and system reliability metrics. Moreover, the multistage robust UC model with dynamic uncertainty sets and policy-guided look-ahead ED dominates RobUC-Static. RobUC-Dynamic also exhibits higher utilization of storage devices and reduces renewable curtailment.

4.6 Conclusion

We present a multistage robust UC model with dynamic uncertainty sets for power systems with significant wind and solar power and storage units. We also propose a novel dispatch process to accompany the robust UC model. An efficient solution framework based on constraint generation and duality reformulations, with several algorithmic improvements, is developed. With extensive computational experiments,

we show that the proposed algorithm can solve large-scale multistage robust UC models with high dimensional uncertainty in a time budget suitable for the day-ahead operation. The proposed robust UC model with the novel ED method is shown to dominate the deterministic UC with reserve and look-ahead ED in both operational cost and system reliability. The proposed dynamic uncertainty sets also effectively capture the temporal and spatial correlations of wind and solar power, which is important for further improving the performance of the multistage robust UC model. The new ED method leads to more utilization of storage units and less curtailment of renewable power.

In summary, the proposed multistage robust UC model, the dynamic uncertainty sets, the policy-guided look-ahead ED, and the solution methodology significantly improve over the existing deterministic and multistage robust UC models and solution methods, and provide a novel and effective approach for operating large-scale power systems with a large number of wind and solar farms and storage devices. Finally, a challenging topic for future work is to incorporate security constraints into the multistage robust UC framework.

CHAPTER V

THE ADAPTIVE ROBUST MULTI-PERIOD ALTERNATING CURRENT OPTIMAL POWER FLOW PROBLEM

5.1 Introduction

The effective operation of power systems is significantly supported by key optimization-based methodologies. A fundamental optimization problem underlying this process is the alternating current (AC) optimal power flow (OPF) problem. This problem consists of minimizing power dispatch costs subject to satisfying power demands and several technical constraints on generating units and the transmission network [36]. Due to the non-convexity of the power flow equations, solving AC OPF to global optimality is tremendously difficult. Given this, and due to the ease of working with linear models, current industry practice involves solving the direct current (DC) approximation of AC OPF, and then adapting the solution obtained to find a feasible AC OPF solution to be implemented [33]. This practice has proven its effectiveness, however, there are significant potential economic and reliability advantages in trying to systematically identify better solutions to AC OPF.

To directly solve AC OPF, there has been important developments in interior point methods that obtain local optimal solutions (see [100] and references therein). The most popular implementation of these methods is MATPOWER [120]. Another important recent line of work is that of convex relaxations of AC OPF. The idea is to develop relaxations of this non-convex problem that can be solved to (near) global optimality. This serves the purpose of bounding the quality of a feasible AC OPF solution and of potentially finding good starting conditions for a local AC OPF solver.

The tightness and efficiency of solving such a relaxation will determine its usefulness. Several relaxations have been proposed and analyzed, including second-order cone programming (SOCP) relaxations [49], [50], [58], semidefinite programming (SDP) relaxations [3], [62], [59], and moment-based relaxations [73]. See [70] for a thorough survey.

Besides trying to improve AC OPF solutions, a challenge of uppermost relevance in today's power systems is their economic and reliable operation under a deep adoption of intermittent renewable power, namely wind and solar power [110]. The push for sustainability has led to an unprecedented increase in the levels of uncertainty under which power systems operate, and new operational and market methodologies are needed to effectively drive an economic and reliable power system under such large adoptions of wind and solar power. In this context, robust optimization has seen a wide reception by industry and academia as a key philosophy that could effectively help meet this challenge. An important example is given by the two-stage adaptive robust optimization models first developed for the day-ahead unit commitment problem in [13], [53], [118]. In these models, there is a set of first-stage decisions that are made before the uncertainty is revealed and encompass a schedule of on/off generator decisions for the next day, and a set of second-stage decisions composed of power dispatch actions that can adapt to the specific realization of uncertainty. After these works, a wide literature on robust optimization for power system operations has started to develop. Other important works on robust unit commitment have considered generator and transmission line contingencies [94], [103], uncertainty in wind power and demand response [116], the use of approximations of the AC OPF equations [1], and multistage models with affine policies [69], [68], to name a few works.

Robust optimization has also seen the recent development of models for real-time operations. [111] integrated statistical methods for wind forecasting with a

static robust look-ahead economic dispatch; [51], [66] and [52] employed affine policies for OPF; [109] developed a robust look-ahead economic dispatch model based on allowable intervals for wind power generation; [117] studied the problem of identifying the largest operating ranges of variable resources that maintain the feasibility of the system; and [67] proposed a two-stage economic dispatch model and the concept of dynamic uncertainty sets for wind speeds. All of the above mentioned works on robust optimization for power system operations have considered DC OPF models, with the exception of [1], where AC OPF equations are approximated using Taylor expansions and a "signomial transformation".

Other methodologies from optimization under uncertainty have also been applied for OPF problems. [105], [84], [92] present stochastic optimization methods for OPF. [105] uses affine policies and considers storage units, under a DC model; [84] considers an interesting two-stage stochastic AC OPF model, using a sample-average approximation approach; and [92] proposes a two-stage stochastic security-constrained DC OPF model under uncertainty in transmission line contingencies. Other related methods that have been employed for DC OPF are distributionally robust optimization [65], [95] and chance constraints [15], [71].

This Chapter addresses the challenge of jointly considering the inherent non-convexity in the AC OPF equations and systematically modeling uncertainty through robust optimization. The contributions of this Chapter are summarized as follows:

1. A two-stage adaptive robust optimization model for the AC OPF problem under uncertainty in active and reactive power demand injections and the availability of renewable power is developed. The model includes detailed technical aspects such as the reactive capability curves of conventional generators and renewable units.

2. To solve this model, we propose to make use of convex relaxations of AC OPF combined with an alternating direction method to find worst-case uncertainty

realizations. We also develop a transmission line screening method to speed-up the algorithm.

3. In extensive computational experiments we explore the performance of different approximations of AC OPF to find worst-case uncertainty realizations, show that the model proposed can be efficiently solved, and analyze the potential practical benefits of the proposed approach using a realistic simulation platform.

The remainder of this Chapter is organized as follows. Section 5.2 reviews known deterministic OPF models and formulates the Robust AC OPF problem. Section 5.3 proposes solution methods for this problem. Section 5.4 presents extensive computational experiments to understand how the different approximations of AC OPF can help in finding worst-case uncertainty realizations, the efficiency of the solution method, and the practical benefits of the proposed approach. Section 5.5 concludes the Chapter.

5.2 Mathematical Models

5.2.1 Deterministic OPF models

Consider the multi-period AC OPF problem (denoted AC-OPF):

$$\min_{\boldsymbol{p}^g,\, \boldsymbol{q}^g,\, \boldsymbol{c},\, \boldsymbol{s},\, \boldsymbol{\theta}} \quad \sum_{i\in\mathcal{G}} \sum_{t\in\mathcal{T}} C^g_{it}\, p^g_{it} \tag{57a}$$

$$\text{s.t.} \quad \underline{p}^g_{it} \le p^g_{it} \le \overline{p}^g_{it} \quad \forall\, i \in \mathcal{G},\, t \in \mathcal{T} \tag{57b}$$

$$\underline{q}^g_{it} \le q^g_{it} \le \overline{q}^g_{it} \quad \forall\, i \in \mathcal{G},\, t \in \mathcal{T} \tag{57c}$$

$$-RD_{it} \le p^g_{it} - p^g_{i,t-1} \le RU_{it} \quad \forall\, i \in \mathcal{G},\, t \in \mathcal{T} \tag{57d}$$

$$a^g_{im}\, p^g_{it} + b^g_{im}\, q^g_{it} \le h^g_{im} \quad \forall\, i \in \mathcal{G},\, t \in \mathcal{T},\, m \in \mathcal{M} \tag{57e}$$

$$\underline{V}_i^2 \le c_{iit} \le \overline{V}_i^2 \quad \forall\, i \in \mathcal{B},\, t \in \mathcal{T} \tag{57f}$$

$$\sum_{k\in\mathcal{G}(i)} p^g_{kt} - p^d_{it}$$

$$= \sum_{j \in \delta(i)} (G_{ij}\, c_{ijt} - B_{ij}\, s_{ijt}) \quad \forall i \in \mathcal{B}, t \in \mathcal{T} \tag{57g}$$

$$\sum_{k \in \mathcal{G}(i)} q^g_{kt} - q^d_{it}$$

$$= \sum_{j \in \delta(i)} (-B_{ij}\, c_{ijt} - G_{ij}\, s_{ijt}) \quad \forall i \in \mathcal{B}, t \in \mathcal{T} \tag{57h}$$

$$(-G_{ij}\, c_{iit} + G_{ij}\, c_{ijt} - B_{ij}\, s_{ijt})^2$$

$$+ (B_{ij}\, c_{iit} - B_{ij}\, c_{ijt} - G_{ij}\, s_{ijt})^2 \leq (f^{max}_{ij})^2$$

$$\forall\, (i,j) \in \mathcal{L},\, t \in \mathcal{T} \tag{57i}$$

$$c_{ijt} = c_{jit}, \quad s_{ijt} = -s_{jit} \quad \forall\, (i,j) \in \mathcal{L},\, t \in \mathcal{T} \tag{57j}$$

$$c^2_{ijt} + s^2_{ijt} = c_{iit}\, c_{jjt} \quad \forall\, (i,j) \in \mathcal{L},\, t \in \mathcal{T} \tag{57k}$$

$$\theta_{jt} - \theta_{it} = \text{arctg}\,(s_{ijt}/c_{ijt}) \quad \forall\, (i,j) \in \mathcal{L},\, t \in \mathcal{T}. \tag{57l}$$

Here, $\mathcal{B}, \mathcal{G}, \mathcal{L}, \mathcal{T}$ are the set of buses, generators, transmission lines, and time periods, respectively. Decisions p^g_{it} and q^g_{it} are the active and reactive power output of generator i at time t, respectively, and θ_{it} is the voltage angle at bus i, time t. The other decisions correspond to $c_{ijt} = |V_{it}||V_{jt}| \cos(\theta_{it} - \theta_{jt})$ and $s_{ijt} = -|V_{it}||V_{jt}| \sin(\theta_{it} - \theta_{jt})$, for buses i, j at time t, where $|V_{it}|$ is the voltage magnitude at bus i, time t. For generator i at time t, parameters C^g_{it}, $\underline{p}^g_{it}$, $\overline{p}^g_{it}$, $\underline{q}^g_{it}$, $\overline{q}^g_{it}$, RD_{it}, RU_{it} correspond to the unit cost, lower and upper bound on active power output, lower and upper bound on reactive power output, and ramp-down and ramp-up capacities, respectively. Parameters a^g_{im}, b^g_{im}, h^g_{im} determine a linear constraint for active and reactive power output at generator i. For bus i, parameters $\underline{V}_i$, $\overline{V}_i$ are the lower and upper bounds on voltage magnitude, respectively, $\mathcal{G}(i)$ is the set of generators at this bus, and $\delta(i)$ is the set of buses connected through a transmission line. For buses i, j, parameters G_{ij} and B_{ij} are the real and imaginary parts of the corresponding admittance matrix element, respectively, and f^{max}_{ij} is the maximum apparent power flow allowed on transmission line (i, j).

In AC-OPF (57), objective (57a) represents total dispatch cost minimization. Constraints (57b)-(57c) impose active and reactive power output bounds; (57d) enforces ramping capacities; and (57e) represents a polyhedral model for reactive capability curve limitations. Constraints (57f) are voltage magnitude bounds; (57g)-(57h) are the power flow equations that determine active and reactive power balance; and (57i) represents transmission line power flow capacities. Finally, constraints (57j)-(57l) are required relations between the $\boldsymbol{c}, \boldsymbol{s}, \boldsymbol{\theta}$ variables that guarantee the consistency of the formulation (see [58] for details on formulating AC OPF using the $\boldsymbol{c}, \boldsymbol{s}$ variables).

The AC OPF problem can be equivalently formulated as a quadratically constrained optimization problem through the following "rectangular" formulation (denoted EF-AC-OPF):

$$\min_{\boldsymbol{p}^g, \boldsymbol{q}^g, \boldsymbol{c}, \boldsymbol{s}, \boldsymbol{e}, \boldsymbol{f}} \quad \sum_{i \in \mathcal{G}} \sum_{t \in \mathcal{T}} C^g_{it}\, p^g_{it} \tag{58a}$$

$$\text{s.t.} \quad \text{eqs. (57b)-(57i) hold} \tag{58b}$$

$$c_{ijt} = e_{it}\, e_{jt} + f_{it}\, f_{jt} \quad \forall\, (i,j) \in \mathcal{L},\, t \in \mathcal{T} \tag{58c}$$

$$s_{ijt} = e_{it}\, f_{jt} - f_{it}\, e_{jt} \quad \forall\, (i,j) \in \mathcal{L},\, t \in \mathcal{T}. \tag{58d}$$

Here $\boldsymbol{e}, \boldsymbol{f}$ are such that the voltage phasor at bus i and time t is given by $V_{it} = e_{it} + \mathrm{i}\, f_{it}$.

As seen from formulations (57) and (58), AC OPF is a non-convex optimization problem. Given the intrinsic difficulty this poses, it is useful to consider *convex relaxations* that can be solved to global optimality in polynomial time. They serve the purpose of providing an approximation for AC OPF, a guaranteed lower bound, and potentially a good starting point for AC OPF algorithms. In this Chapter, we will consider two such relaxations. First, the SOCP relaxation of AC-OPF (57) (denoted SOCP-OPF) (see [42], [50], [58]) is given by:

$$\min_{\boldsymbol{p}^g, \boldsymbol{q}^g, \boldsymbol{c}, \boldsymbol{s}} \quad \sum_{i \in \mathcal{G}} \sum_{t \in \mathcal{T}} C^g_{it}\, p^g_{it} \tag{59a}$$

$$\text{s.t.} \quad \text{eqs. (57b)-(57j) hold} \tag{59b}$$

$$c_{ijt}^2 + s_{ijt}^2 \leq c_{iit}\, c_{jjt} \quad \forall\, (i,j) \in \mathcal{L},\, t \in \mathcal{T}. \tag{59c}$$

Second, the SDP relaxation of EF-AC-OPF (58) (denoted SDP-OPF) (see [3], [58]) is given by:

$$\min_{\boldsymbol{p}^g,\, \boldsymbol{q}^g,\, \boldsymbol{c},\, \boldsymbol{s},\, \boldsymbol{X}} \quad \sum_{i \in \mathcal{G}} \sum_{t \in \mathcal{T}} C_{it}^g\, p_{it}^g \tag{60a}$$

$$\text{s.t.} \quad \text{eqs. (57b)-(57i) hold} \tag{60b}$$

$$c_{ijt} = X_{ijt} + X_{i'j't} \quad \forall\, (i,j) \in \mathcal{L},\, t \in \mathcal{T} \tag{60c}$$

$$s_{ijt} = X_{ij't} - X_{i'jt} \quad \forall\, (i,j) \in \mathcal{L},\, t \in \mathcal{T} \tag{60d}$$

$$\boldsymbol{X}_t \succeq 0 \quad \forall\, t \in \mathcal{T}, \tag{60e}$$

where, $i' = i + |\mathcal{B}|$, $j' = j + |\mathcal{B}|$, and $\boldsymbol{X}_t$ is a symmetric positive-semidefinite $2|\mathcal{B}| \times 2|\mathcal{B}|$ matrix.

Besides relaxations of AC OPF, we will also consider the widely used "direct current approximation" of AC-OPF (denoted DC-OPF), given by:

$$\min_{\boldsymbol{p}^g,\, \boldsymbol{c},\, \boldsymbol{s},\, \boldsymbol{\theta}} \quad \sum_{i \in \mathcal{G}} \sum_{t \in \mathcal{T}} C_{it}^g\, p_{it}^g \tag{61a}$$

$$\text{s.t.} \quad \text{eqs. (57b), (57d), (57g) hold} \tag{61b}$$

$$-B_{ij}\, s_{ijt} \leq f_l^{max} \quad \forall\, (i,j) \in \mathcal{L},\, t \in \mathcal{T} \tag{61c}$$

$$c_{ijt} = 1, \quad s_{ijt} = \theta_{jt} - \theta_{it} \quad \forall\, (i,j) \in \mathcal{L},\, t \in \mathcal{T}. \tag{61d}$$

The purpose of having multiple time periods in the deterministic models above is to account for the dynamics of the power system in future time periods. That is, the model is repeatedly run implementing only the decisions corresponding to the first time period, while the decisions corresponding to future time periods are discarded. In this way, the model *looks-ahead* at the system in order to account for the effects that presently implemented decisions will have on the system in the future. A critical limitation of these deterministic models is that they do not account for the

uncertainty in several parameters that affect these future decisions. In particular, a critical data input to a multi-period OPF is the forecast considered for uncertain parameters. In contrast, in what follows we will present a systematic approach for modeling uncertainty in OPF.

5.2.2 Robust OPF models

We can now describe the Robust OPF framework. Consider a time horizon encompassing time periods $t = 0, \ldots, T$. We will assume that in this horizon, all parameters are exactly known at $t = 0$ while there are uncertain parameters at $t = 1, \ldots, T$. We will formulate an optimization problem where dispatch decisions at $t = 0$ are selected in such a way that the worst-case costs are minimized over $t = 0, \ldots, T$, given that uncertain parameters lie in some uncertainty set and dispatch decisions at $t = 1, \ldots, T$ can *adapt* to the realization of these uncertain parameters.

The general Robust OPF problem is formulated in compact form as follows:

$$\min_{(\boldsymbol{p}_0^g, \boldsymbol{q}_0^g) \in \Omega_0} \Big(\boldsymbol{C}_0^\top \boldsymbol{p}_0^g + \max_{(\boldsymbol{p}^d, \boldsymbol{q}^d) \in \mathcal{D}, \overline{\boldsymbol{p}}^r \in \overline{\mathcal{P}}^r} \quad \min_{(\boldsymbol{p}^g, \boldsymbol{q}^g) \in \Omega(\boldsymbol{p}_0^g, \boldsymbol{p}^d, \boldsymbol{q}^d, \overline{\boldsymbol{p}}^r)} \sum_{t \in \mathcal{T}} \boldsymbol{C}_t^\top \boldsymbol{p}_t^g \Big). \tag{62}$$

Here Ω_0 represents all OPF constraints at $t = 0$, $\mathcal{D}$ is an uncertainty set for active ($\boldsymbol{p}^d$) and reactive ($\boldsymbol{q}^d$) power demands over $t = 1, \ldots, T$, $\overline{\mathcal{P}}^r$ is an uncertainty set for available power ($\overline{\boldsymbol{p}}^r$) at all renewable units over $t = 1, \ldots, T$, and $\Omega(\boldsymbol{p}_0^g, \boldsymbol{p}^d, \boldsymbol{q}^d, \overline{\boldsymbol{p}}^r)$ represents all OPF constraints over $t = 1, \ldots, T$ given dispatch decisions at $t = 0$ as well as realized active and reactive power demands and available power at all renewable units over $t = 1, \ldots, T$.

In (62), we assume that there is a set $\mathcal{R}$ of renewable units that is a subset of the set of generators $\mathcal{G}$ and we have $\overline{p}_{it}^g = \overline{p}_{it}^r$ for $i \in \mathcal{R}$. With this notation, we do not need to separately write equations for wind and solar farms.

The problem structure of the general Robust OPF (62) is flexible. Depending on the OPF formulation considered, a different problem is obtained. In particular,

we define the Robust AC-OPF problem as problem (62) under AC-OPF formulation (57), that is, problem (62) with

$$\Omega_0 = \Big\{(\boldsymbol{p}_0^g, \boldsymbol{q}_0^g) : \exists\, \boldsymbol{c}_0, \boldsymbol{s}_0, \boldsymbol{\theta}_0 \\ \text{s.t. eqs. (57b)-(57l) hold under } \mathcal{T} = \{0\}\Big\},$$

$$\Omega(\boldsymbol{p}_0^g, \boldsymbol{p}^d, \boldsymbol{q}^d, \overline{\boldsymbol{p}}^r) = \Big\{(\boldsymbol{p}^g, \boldsymbol{q}^g) : \exists\, \boldsymbol{c}, \boldsymbol{s}, \boldsymbol{\theta} \\ \text{s.t. eqs. (57b)-(57l) hold under } \mathcal{T} = \{1, \ldots, T\}\Big\}.$$

Similarly, one can define the Robust SOCP-OPF (with (59)), Robust SDP-OPF (with (60)) and Robust DC-OPF (with (61)).

5.2.3 The conic cases and the use of conic duality

SOCP-, SDP- and DC-OPF share a common structure. These are *conic* optimization problems [8]. These problems are based on the Lorentz, semidefinite, and nonnegative orthant cones, respectively. Let's write problem (62) with shorter notation as

$$\min_{\boldsymbol{x} \in X} \left(\boldsymbol{C}^\top \boldsymbol{x} + \max_{\boldsymbol{\xi} \in \Xi} \min_{\boldsymbol{y} \in Y(\boldsymbol{x}, \boldsymbol{\xi})} \boldsymbol{b}^\top \boldsymbol{y} \right), \tag{63}$$

where $X = \Omega_0$, $\Xi = \mathcal{D} \times \overline{\mathcal{P}}^r$ and $Y(\boldsymbol{x}, \boldsymbol{\xi}) = \Omega(\boldsymbol{p}_0^g, \boldsymbol{p}^d, \boldsymbol{q}^d, \overline{\boldsymbol{p}}^r)$. Now, under conic OPF formulations, we have $Y(\boldsymbol{x}, \boldsymbol{\xi}) = \{\boldsymbol{y} : \boldsymbol{A}\boldsymbol{y} \geq_K \boldsymbol{f} + \boldsymbol{G}\boldsymbol{x} + \boldsymbol{H}\boldsymbol{\xi}\}$, where K is a closed, pointed and self-dual cone with non-empty interior. Given this, we can exploit conic duality to reformulate the problem. Assuming $Y(\boldsymbol{x}, \boldsymbol{\xi})$ is bounded with non-empty interior, strong duality of the inner min problem holds [8], which leads to

$$\max_{\boldsymbol{\xi} \in \Xi} \min_{\boldsymbol{y}} \Big\{ \boldsymbol{b}^\top \boldsymbol{y} : \boldsymbol{A}\,\boldsymbol{y} \geq_K \boldsymbol{f} + \boldsymbol{G}\,\boldsymbol{x} + \boldsymbol{H}\boldsymbol{\xi} \Big\} \tag{64a}$$

$$= \max_{\boldsymbol{\xi} \in \Xi} \max_{\boldsymbol{\pi}} \Big\{ \boldsymbol{\pi}^\top (\boldsymbol{f} + \boldsymbol{G}\boldsymbol{x} + \boldsymbol{H}\boldsymbol{\xi}) : \boldsymbol{\pi}^\top \boldsymbol{A} = \boldsymbol{b}^\top, \boldsymbol{\pi} \geq_K 0 \Big\} \tag{64b}$$

$$= \max_{\boldsymbol{\xi}, \boldsymbol{\pi}} \Big\{ \boldsymbol{\pi}^\top (\boldsymbol{f} + \boldsymbol{G}\boldsymbol{x}) + \boldsymbol{\pi}^\top \boldsymbol{H}\boldsymbol{\xi} : \boldsymbol{\xi} \in \Xi, \boldsymbol{\pi}^\top \boldsymbol{A} = \boldsymbol{b}^\top, \boldsymbol{\pi} \geq_K 0 \Big\}. \tag{64c}$$

With this, and assuming Ξ is a conic set (i.e. $\Xi = \{\boldsymbol{\xi} : W\boldsymbol{\xi} \geq_{\hat{K}} \boldsymbol{w}\}$ for some closed, pointed cone with non-empty interior $\hat{K}$), the max-min problem (64a) can be reformulated as a bi-conic optimization problem (64c). In particular, if the uncertainty

set Ξ is a polytope and the cone K is the non-negative orthant (DC-OPF case) then the max-min problem reduces to a bilinear optimization problem.

5.2.4 Uncertainty set

In this Chapter, we consider the following "budget" uncertainty set for active and reactive power demands:

$$\mathcal{D} = \Big\{(\boldsymbol{p}^d, \boldsymbol{q}^d) : \sum_{i\in\mathcal{B}} \left|p^d_{it} - \tilde{p}^d_{it}\right| / \hat{p}^d_{it} \le \Gamma\sqrt{|\mathcal{B}|} \tag{65a}$$

$$\sum_{i\in\mathcal{B}} \left|q^d_{it} - \tilde{q}^d_{it}\right| / \hat{q}^d_{it} \le \Gamma\sqrt{|\mathcal{B}|} \tag{65b}$$

$$p^d_{it} \in [\tilde{p}^d_{it} - \Gamma\hat{p}^d_{it}, \tilde{p}^d_{it} + \Gamma\hat{p}^d_{it}]\, \forall i \in \mathcal{B}, t \in \mathcal{T} \tag{65c}$$

$$q^d_{it} \in [\tilde{q}^d_{it} - \Gamma\hat{q}^d_{it}, \tilde{q}^d_{it} + \Gamma\hat{q}^d_{it}]\, \forall i \in \mathcal{B}, t \in \mathcal{T}\Big\}. \tag{65d}$$

Here, $\tilde{p}^d_{it}$, $\tilde{q}^d_{it}$ are the nominal values for active and reactive power demand at bus i, time t, respectively, and $\hat{p}^d_{it}$, $\hat{q}^d_{it}$ are the respective variability factors (e.g. standard deviation from historic data). Parameter Γ determines the size of the uncertainty set.

For renewable power, we consider the following "dynamic" uncertainty set, recently proposed in [68]:

$$\overline{\mathcal{P}}^r = \Big\{\overline{\boldsymbol{p}}^r = (\overline{p}^r_{it})_{it} \,:\, \exists\, \boldsymbol{u},\, \boldsymbol{v} \;\text{ s.t.} \tag{66a}$$

$$\overline{p}^r_{it} = f_{it} + g_{it}\, u_{it} \quad \forall\, i \in \mathcal{R},\, t \in \mathcal{T} \tag{66b}$$

$$\boldsymbol{u}_t = \sum_{l=1}^{L} \boldsymbol{A}^l\, \boldsymbol{u}_{t-l} + \boldsymbol{B}\, \boldsymbol{v}_t \quad \forall\, t \in \mathcal{T} \tag{66c}$$

$$\|\boldsymbol{v}_t\| \le \Gamma \quad \forall\, t \in \mathcal{T} \tag{66d}$$

$$\sum_{t\in\mathcal{T}} \|\boldsymbol{v}_t\| \le \rho\, \Gamma\, |\mathcal{T}| \tag{66e}$$

$$0 \le \overline{p}^r_{it} \le \overline{p}^{r,max}_{it} \quad \forall\, i \in \mathcal{R},\, t \in \mathcal{T}\Big\}. \tag{66f}$$

Here, $\overline{p}^r_{it}$ is the available renewable power at renewable unit i at time t, parameters f_{it} and g_{it} are deterministic seasonality components, matrices $\boldsymbol{A}^l$ and $\boldsymbol{B}$ determine

spatial and temporal correlations (of up to L time periods), $\|\cdot\|$ is a norm, Γ is an uncertainty set size parameter, ρ determines a "budget of uncertainty over time periods", and $\overline{p}_{it}^{r,max}$ is an upper bound on $\overline{p}_{it}^{r}$. Component $\boldsymbol{u}_t \in \mathbb{R}^{|\mathcal{R}|}$ can be seen as a residual uncertainty after seasonal components are filtered out, and $\boldsymbol{v}_t \in \mathbb{R}^{N_v}$ as a primitive uncertainty for which temporal and spatial correlations are further filtered out as well, where $N_v \in \{1, \ldots, |\mathcal{R}|\}$. See [68] for more details about this dynamic uncertainty set model.

5.3 Solution Methods for Robust OPF

In this Section we propose solution strategies for the Robust OPF problem (63). Section 5.3.1 reviews a general algorithmic framework. Section 5.3.2 studies a method to solve the max-min problem under conic OPF formulations. In Section 5.3.3 we propose a method to solve the max-min problem under AC-OPF. Section 5.3.4 develops an important speed-up technique.

5.3.1 Overall solution framework

The Robust OPF problem (63) is equivalent to

$$\min_{\boldsymbol{x}\in X,\eta} \ \boldsymbol{C}^\top \boldsymbol{x} + \eta \tag{67a}$$

$$\text{s.t.} \quad \eta \geq \min_{y\in Y(\boldsymbol{x},\boldsymbol{\xi})} \boldsymbol{b}^\top \boldsymbol{y} \quad \forall \boldsymbol{\xi} \in \Xi, \tag{67b}$$

which can also be formulated as

$$\min_{\boldsymbol{x}\in X,\eta,\boldsymbol{y}(\cdot)} \ \boldsymbol{C}^\top \boldsymbol{x} + \eta \tag{68a}$$

$$\text{s.t.} \quad \eta \geq \boldsymbol{b}^\top \boldsymbol{y}(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \in \Xi \tag{68b}$$

$$\boldsymbol{y}(\boldsymbol{\xi}) \in Y(\boldsymbol{x}, \boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \in \Xi. \tag{68c}$$

This problem structure suggests an iterative solution method [13], [113]. The idea is to repeatedly take a solution $\boldsymbol{x}$ and solve

$$Q(\boldsymbol{x}) = \max_{\boldsymbol{\xi}\in\Xi} \min_{\boldsymbol{y}\in Y(\boldsymbol{x},\boldsymbol{\xi})} \boldsymbol{b}^\top \boldsymbol{y}, \tag{69a}$$

to then solve the following master problem:

$$\min_{\boldsymbol{x}\in X,\eta\geq 0,\boldsymbol{y}(\cdot)} \boldsymbol{C}^\top \boldsymbol{x} + \eta \tag{70a}$$

$$\text{s.t.} \quad \eta \geq \boldsymbol{b}^\top \boldsymbol{y}(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \in \{\boldsymbol{\xi}_1^*, \ldots, \boldsymbol{\xi}_k^*\} \tag{70b}$$

$$\boldsymbol{y}(\boldsymbol{\xi}) \in Y(\boldsymbol{x}, \boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \in \{\boldsymbol{\xi}_1^*, \ldots, \boldsymbol{\xi}_k^*\}, \tag{70c}$$

where $\boldsymbol{\xi}_k^*$ is the solution found to (69) in the k-th iteration. This solution method is formally presented in Algorithm 6.

Algorithm 6 Solution method for Robust OPF (63)

1: $k = 0$
2: **repeat**
3: $(\boldsymbol{x}, \eta) \leftarrow$ optimal solution of (70)
4: $k \leftarrow k + 1$
5: Evaluate $Q(\boldsymbol{x})$: $\boldsymbol{\xi}_k^* \leftarrow$ optimal solution of (69)
6: **until** $Q(\boldsymbol{x}) \leq \eta$

In [113] the authors show that Algorithm 6 converges in a finite number of iterations to the optimal solution of (63) when Ξ is a polytope and $Y(\boldsymbol{x}, \boldsymbol{\xi})$ has a polyhedral structure given by $Y(\boldsymbol{x}, \boldsymbol{\xi}) = \{\boldsymbol{y} : \boldsymbol{A}\boldsymbol{y} \geq \boldsymbol{f} + \boldsymbol{G}\boldsymbol{x} + \boldsymbol{H}\boldsymbol{\xi}\}$. For other cases this does not necessarily hold, but the method can be used to find good solutions.

A difficult step in Algorithm 6 is evaluating $Q(\boldsymbol{x})$. We will study this in what follows.

5.3.2 An alternating direction method for conic OPF cases

As shown in eq. (64), under conic structures given by $Y(\boldsymbol{x}, \boldsymbol{\xi}) = \{\boldsymbol{y} : \boldsymbol{A}\boldsymbol{y} \geq_K \boldsymbol{f} + \boldsymbol{G}\boldsymbol{x} + \boldsymbol{H}\boldsymbol{\xi}\}$ and $\Xi = \{\boldsymbol{\xi} : W\boldsymbol{\xi} \geq_{\hat{K}} w\}$ evaluating $Q(\boldsymbol{x})$ accounts to solving a bi-conic optimization problem (64c). This means that if $\boldsymbol{\xi}$ is fixed problem (64c) becomes a conic problem:

$$\max_{\boldsymbol{\pi}} \left\{ \boldsymbol{\pi}^\top (\boldsymbol{f} + \boldsymbol{G}\boldsymbol{x}) + \boldsymbol{\pi}^\top \boldsymbol{H}\boldsymbol{\xi} : \boldsymbol{\pi}^\top \boldsymbol{A} = \boldsymbol{b}^\top, \boldsymbol{\pi} \geq_K 0 \right\}, \tag{71}$$

and if $\boldsymbol{\pi}$ is fixed problem (64c) also becomes a conic problem:

$$\max_{\boldsymbol{\xi}} \left\{ \boldsymbol{\pi}^\top (\boldsymbol{f} + \boldsymbol{G}\boldsymbol{x}) + \boldsymbol{\pi}^\top \boldsymbol{H}\boldsymbol{\xi} : \boldsymbol{\xi} \in \Xi \right\}. \tag{72}$$

This suggests an iterative method, first proposed in [60] (for the bilinear case, but with a natural extension to the bi-conic case), which we call "alternating direction". The idea is to fix some starting $\boldsymbol{\xi}$ and optimize over $\boldsymbol{\pi}$, to then fix the obtained $\boldsymbol{\pi}$ and optimize over $\boldsymbol{\xi}$, and so on until convergence. This method is formally presented in Algorithm 7.

Algorithm 7 Alternating direction method for (64c)

1: Choose an initial $\boldsymbol{\xi} \in \Xi$
2: **repeat**
3: $\quad \boldsymbol{\pi} \leftarrow$ optimal solution of (71) with objective θ
4: $\quad \boldsymbol{\xi} \leftarrow$ optimal solution of (72) with objective θ'
5: **until** $\theta' = \theta$
6: Output: θ as estimate of $Q(\boldsymbol{x})$ with solution $\boldsymbol{\xi}^*$

In the bilinear case, Algorithm 7 converges monotonically to a KKT point of problem (64c) in a finite number of iterations [60]. In the general bi-conic case only monotonic convergence is ensured, however, in practice the method can be stopped when θ' is close to θ.

5.3.3 Using conic approximations in the AC-OPF case

When $Y(\boldsymbol{x}, \boldsymbol{\xi})$ represents AC-OPF (57) we cannot reformulate the max-min problem defining $Q(\boldsymbol{x})$ in (69). Given this, we propose to estimate $Q(\boldsymbol{x})$ by first approximating AC-OPF through SOCP-, SDP- or DC-OPF, running the corresponding alternating direction method (Algorithm 7) to obtain a solution $\boldsymbol{\xi}^*$, and then solving AC-OPF, taking $\min\{\boldsymbol{b}^\top \boldsymbol{y} : \boldsymbol{y} \in Y(\boldsymbol{x}, \boldsymbol{\xi}^*)\}$ as estimate of $Q(\boldsymbol{x})$.

5.3.4 Screening transmission line constraints

We describe here a critical enhancement for accelerating Algorithm 6 (under any OPF formulation). Usually, power systems have many transmission lines among which only a subset of them is close to or at full power flow capacity. Therefore, we propose to augment Algorithm 6 by iteratively adding transmission line constraints on an as-needed basis. The idea is to start with a subset of these constraints in both the

master problem (70) and the max-min problem (69) and in every iteration check if the constraints not included are satisfied or not. If they are not satisfied, they are incorporated to the problem in explicit form. The resulting enhanced solution method is presented in Algorithm 8.

Algorithm 8 Enhanced solution method for Robust OPF (63)

Choose initial sets $\mathcal{L}_M$, $\mathcal{L}_S$ of transmission constraints (57i) to consider in the master (70) and max-min (69) problems
$k = 0$
repeat
  $(\boldsymbol{x}, \eta) \leftarrow$ optimal solution of (70) under $\mathcal{L}_M$
  for all $(i, j) \in \mathcal{L}$ **do**
    If the (i, j)-th transmission constraints are not satisfied in (70), do $\mathcal{L}_M \leftarrow \mathcal{L}_M \cup \{(i, j)\}$
  end for
  $k \leftarrow k + 1$
  repeat
    Evaluate $Q(\boldsymbol{x})$: $\boldsymbol{\xi}_k^* \leftarrow$ optimal solution of (69) under $\mathcal{L}_S$
    for all $(i, j) \in \mathcal{L}$ **do**
      If the (i, j)-th transmission constraints are not satisfied in (69), do $\mathcal{L}_S \leftarrow \mathcal{L}_S \cup \{(i, j)\}$
    end for
  until All transmission constraints in (69) are satisfied
until $Q(\boldsymbol{x}) \leq \eta$ and all transmission constraints are satisfied

5.4 Computational Experiments

We conduct extensive computational experiments to study the performance of the algorithms developed as well as the potential advantages in practice of the Robust AC-OPF. To carry out these experiments, we work with adapted versions of the IEEE 14- 118-, and 2736-bus systems [120]. The main features of the test cases employed are summarized in Table 20. Reactive capability curves of conventional generators, wind farms and solar farms are adapted from [32]. All experiments have been performed using Python 2.7 in a PC with an Intel Core i5 processor at 2.4 GHz with 4GB memory, using GUROBI as linear and SOCP solver, MOSEK as SDP solver and IPOPT as local solver for AC-OPF master- and sub-problems. In the

Table 20: Main features of the test cases employed

Buses	14	118	2736
Conventional generators	5	54	253
Wind farms	4	40	60
Solar farms	1	20	30
Loads	12	99	2011
Lines	20	186	3504
Conventional generation capacity (MW)	400	9966	28868
Installed wind power capacity (MW)	240	2394	8400
Installed solar power capacity (MW)	29	1462	4363
Average total nominal demand (MW)	254	6256	14264
Max total nominal demand (MW)	320	7878	18075

Robust OPF, we consider 4 time periods (one corresponding to first-stage decisions and three corresponding to second-stage decisions). For the dynamic uncertainty set (66) we estimate $\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{A}_l, \boldsymbol{B}$ from data and use $L = 1$, $\rho = 0.2$, several values for the uncertainty set size parameter Γ, and $\|\cdot\| = \max\{\|\cdot\|_1/\sqrt{N_v}, \|\cdot\|_\infty\}$, with which the uncertainty set becomes polyhedral. Section 5.4.1 analyzes the solutions obtained for the max-min component of the Robust OPF. Section 5.4.2 studies the efficiency of the solution method developed for Robust AC-OPF. Section 5.4.3 studies the performance of the Robust AC-OPF in a simulation platform.

5.4.1 Solving the max-min problem

5.4.1.1 Conic OPF formulations

We study the alternating direction method proposed in Section 5.3.2 to solve the max-min problem (69) using DC-, SOCP- and SDP-OPF under different conditions corresponding to a simulation of the power system over 5 days, encompassing a total of 120 problems solved, using "small" ($\Gamma = 0.5$) and "large" ($\Gamma = 2$) uncertainty set sizes. Table 21 summarizes the average running time, number of iterations, and improvement in the objective function when starting the algorithm from multiple initial solutions. We can observe that: the AD method runs extremely fast for the DC-OPF case, relatively fast for the SOCP-OPF case, and slower for the SDP-OPF case

(significantly slower for the 118-bus system); the number of iterations is similar for the three cases; and testing different initial solutions does not significantly help (specially for the SOCP and SDP cases). We also remark that in our experiments, SDP-OPF is very unstable for the 118-bus system, for which numeric problems stopped the algorithm for 32% of the SDP-OPF alternating direction runs.

We are interested in trying to understand whether the solutions obtained are close to optimal or not. Unfortunately, even though we can exploit the bi-conic structure of the problem, this is a difficult non-convex problem (more specifically, a maximization problem with a non-concave objective function) and to the best of our knowledge there are no tight bounds available to calculate a reasonable optimality gap. In particular, for the DC-OPF case we tested a state-of-the-art bound [2] based on SDP relaxation of the bilinear components present in the problem, strengthened with McCormick envelopes, without success (obtaining gaps of more than 100%, which we believe to occur mostly due to the looseness of the variable bounds). We also tested the state-of-the-art global solver Baron [90] for the DC-OPF case, with a time limit of 5 minutes. While the gap can improve significantly (but still being far from good) it is important to note that in every single case the best objective obtained by Baron was either the same obtained by the AD method or slightly lower.

5.4.1.2 AC-OPF

We study the method proposed in Section 5.3.3 to solve the max-min problem under AC-OPF. We solve 120 max-min problems under different conditions corresponding to a simulation over 5 days, using the DC-, SOCP- and SDP-OPF approximations of AC-OPF. For each step of the simulation, we run the alternating direction method with each of these approximations to obtain a corresponding $\boldsymbol{\xi}^{DC}$, $\boldsymbol{\xi}^{SOCP}$ or $\boldsymbol{\xi}^{SDP}$, respectively, and then solve AC-OPF under each of these three $\boldsymbol{\xi}$'s. Then, we rank

Table 21: Max-min problem performance under Conic OPF formulations (average over 120 problems)

14-bus system						
OPF type	DC		SOCP		SDP	
Γ	0.5	2	0.5	2	0.5	2
Running time (s)	0.1	0.1	0.3	0.3	5.1	4.0
Iterations	3.0	3.1	3.0	3.1	2.9	3.2
Impr. mul. ini. sols. (%)	0.00	0.07	0.00	0.00	0.00	0.10
118-bus system						
OPF type	DC		SOCP		SDP	
Γ	0.5	2	0.5	2	0.5	2
Running time (s)	0.8	0.8	1.3	1.7	160.2	151.9
Iterations	3.7	4.1	3.0	4.2	2.4	2.9
Impr. mul. ini. sols. (%)	0.64	0.67	0.00	0.10	0.00	0.01

these $\boldsymbol{\xi}$'s according to their AC-OPF cost (which involves solving a deterministic AC-OPF under each of them), and call $\boldsymbol{\xi}^*$ the one that achieves the highest AC-OPF cost, that is, $\boldsymbol{\xi}^*$ is the best solution identified for the max-min problem (69) under AC-OPF. Table 22 summarizes the results obtained. In this Table "$\boldsymbol{\xi}$ obtained vs $\boldsymbol{\xi}^*$" compares the AC-OPF cost under $\boldsymbol{\xi}^{DC}$, $\boldsymbol{\xi}^{SOCP}$ or $\boldsymbol{\xi}^{SDP}$, respectively, to the AC-OPF cost under $\boldsymbol{\xi}^*$. Also, "Gap under $\boldsymbol{\xi}^*$" compares the AC-OPF cost under $\boldsymbol{\xi}^*$ to the cost obtained when solving DC-OPF, SOCP-OPF or SDP-OPF, respectively, under ξ^*. With this, "$\boldsymbol{\xi}$ obtained vs $\boldsymbol{\xi}^*$" measures the quality of the solution identified for the max-min problem (69), while "Gap under $\boldsymbol{\xi}^*$" measures how close are DC-OPF, SOCP-OPF and SDP-OPF as approximations for AC-OPF.

From the results obtained, we can observe that: i) SDP-OPF is closest as approximation to AC-OPF and the best solutions identified for the max-min problem (69) are naturally obtained by using this approximation; ii) using SOCP-OPF has a better performance than DC-OPF and leads to good solutions under $\Gamma = 0.5$ but looses quality under $\Gamma = 2$; and iii) DC-OPF is not a close approximation to AC-OPF as it finds a significantly different cost. From these results, it seems most convenient to use SDP-OPF in the alternating direction method, however, as mentioned in Section

Table 22: Max-min problem performance under AC-OPF (average over 120 problems)

14-bus system						
OPF type	DC		SOCP		SDP	
Γ	0.5	2	0.5	2	0.5	2
$\boldsymbol{\xi}$ obtained vs $\boldsymbol{\xi}^*$ (%)	-0.84	-2.91	-0.13	-2.65	0.00	-0.02
Gap under $\boldsymbol{\xi}^*$ (%)	-57.15	-60.42	3.32	8.40	0.00	0.00
118-bus system						
OPF type	DC		SOCP		SDP	
Γ	0.5	2	0.5	2	0.5	2
$\boldsymbol{\xi}$ obtained vs $\boldsymbol{\xi}^*$ (%)	-0.26	-6.85	-0.12	-2.42	0.00	-1.58
Gap under $\boldsymbol{\xi}^*$ (%)	-41.83	-43.57	2.03	9.59	0.55	5.62

5.4.1.1 this choice is not stable and also very slow. Therefore, we conclude that SOCP-OPF provides a good overall choice as approximation to AC-OPF in the alternating direction method.

5.4.2 Overall algorithmic performance

We now turn our attention to the efficiency of the overall algorithm proposed in Section 5.3.4. We solve 120 Robust AC-OPF problems under different conditions corresponding to a simulation over 5 days, using the SOCP-OPF approximation of AC-OPF in the alternating direction method (a choice we justify from the analysis in Section 5.4.1). Tables 23 and 24 summarize the running time and number of iterations of the algorithm, respectively, with ("With TS") and without ("Without TS") the transmission screening technique described in Section 5.3.4. We can observe that the transmission screening technique can significantly speed-up the solution method while roughly maintaining the average number of iterations. In particular, we obtain a 42% average time reduction for the 118-bus system with $\Gamma = 2$ and a 66% average time reduction for the 2736-bus system with $\Gamma = 3$. Further, we believe that the average running time below 21 seconds obtained for the 118-bus system is very promising for real-time applications of the Robust AC-OPF in medium-sized power systems with a significant adoption of intermittent renewable power.

Table 23: Running time (s) for Robust AC-OPF (average over 120 problems)

14-bus system						
Γ	0.5	1	1.5	2	2.5	3
Without TS	1.0	1.1	1.2	1.2	1.2	1.3
With TS	0.8	0.9	1.0	1.0	1.0	1.0
118-bus system						
Γ	0.5	1	1.5	2	2.5	3
Without TS	10.4	11.6	15.3	25.7	23.1	26.9
With TS	7.7	8.2	11.2	14.9	17.8	20.7
2736-bus system						
Γ	0.5	1	1.5	2	2.5	3
Without TS	798	904	820	936	1144	1343
With TS	342	356	353	394	414	462

Table 24: Number of iterations for Robust AC-OPF (average over 120 problems)

14-bus system						
Γ	0.5	1	1.5	2	2.5	3
Without TS	2.0	2.1	2.2	2.2	2.2	2.2
With TS	2.0	2.1	2.2	2.2	2.2	2.3
118-bus system						
Γ	0.5	1	1.5	2	2.5	3
Without TS	2.0	2.0	2.3	2.4	2.6	2.8
With TS	2.0	2.0	2.3	2.4	2.8	2.9
2736-bus system						
Γ	0.5	1	1.5	2	2.5	3
Without TS	2.0	2.0	2.0	2.1	2.1	2.1
With TS	2.0	2.0	2.0	2.0	2.0	2.1

5.4.3 Model comparison through rolling-horizon simulations

This part compares the practical performance that the Robust AC-OPF can offer, using a simulation platform that mimics the hour to hour operation of the power system under the true full AC-OPF power flow equations. The simulation consists of a rolling-horizon process where Robust AC-OPF is repeatedly run as time moves forward. We consider a simulation horizon of 100 contiguous days (i.e. 2400 hours). The simulation starts by observing uncertain parameters at the first hour and solving the Robust AC-OPF under a horizon that spans from this hour (first-stage decision with known parameters) to three hours ahead (second-stage decisions under uncertain parameters). Then, the decisions for this first hour are implemented and we observe the parameters at the second hour, solving now a Robust AC-OPF that spans this second hour (first-stage decision) and three hours ahead (second-stage decisions). This process is repeated until the last simulated hour, after which we compute several cost and reliability performance metrics achieved by the implemented AC-OPF decisions. To simulate uncertain parameters, we use a normal distribution for active and reactive power demand at each node (assuming temporal and spatial independence), and a vector autoregressive process for wind and solar power [68] (that captures temporal and spatial correlations), for which we use data from NREL's Western Wind and Solar Integration Datasets [85] for parameter estimation. The average renewable penetration under these simulated trajectories is 21.6% for the 14-bus system and 18.8% for the 118-bus system.

To properly measure the economic and reliability performance of the Robust AC-OPF, the model is augmented with penalty variables that allow to shed load at any load bus, with a unit cost of \$5000/MWh.

The results from these simulations are summarized in Table 25, considering several values for the uncertainty set size parameter Γ. In this Table, "Cost Avg" is the average hourly cost over the simulated horizon, "Cost Std" is the standard deviation

Table 25: Rolling-horizon simulation results under Robust AC-OPF

14-bus system						
Γ	0	0.5	1	1.5	2	2.5
Cost Avg ($)	9162	8123	7841	8116	8677	9405
Cost Std ($)	10569	5700	2984	2625	2796	3056
Penalty Avg ($)	1642	540	78	6	0	0
Penalty Freq	4.88%	2.29%	0.50%	0.04%	0.00%	0.00%
T Congestion	4.97%	4.66%	4.31%	3.78%	2.71%	1.71%
Ramp (MW/h)	49	52	57	64	73	79
Renewable Util	99.98%	99.93%	99.75%	99.00%	96.85%	91.45%
118-bus system						
Γ	0	0.5	1	1.5	2	2.5
Cost Avg ($)	181625	155358	149778	153822	163126	181540
Cost Std ($)	206116	102140	59992	63461	68016	68415
Penalty Avg ($)	37032	9717	181	0	0	0
Penalty Freq	7.62%	3.33%	0.36%	0.00%	0.00%	0.00%
T Congestion	2.83%	2.72%	2.43%	2.05%	1.61%	1.07%
Ramp (MW/h)	976	1024	1110	1213	1361	1528
Renewable Util	99.99%	99.99%	99.99%	99.99%	99.98%	99.95%

of hourly cost, "Penalty Avg" is the average hourly penalty cost, "Penalty Freq" is the proportion of time periods where penalty occurred, "T Congestion" is the average transmission congestion, based on a metric corresponding to the proportion of transmission line capacities that have an utilization of 90% or more, "Ramp" is the average system-level ramp-up capacity, and "Renewable Util" is the average utilization of renewable power, corresponding to the proportion of dispatched renewable power with respect to its availability.

We can first observe that the tradeoff between economic effectiveness and system reliability can be controlled through the uncertainty set size parameter Γ. As Γ increases we can see that: average costs tend first to decrease (until $\Gamma = 1$ for both systems) and then to increase, cost standard deviation decreases sharply first and then slowly increases, penalty costs and frequency decrease until being completely eliminated, transmission congestion decreases, ramping capacities increase, and renewable power can be slightly curtailed. These results show that the value of Γ can

be used to "tune" the conservativeness with which the power system is operated. A higher Γ provides a more reliable operation, but potentially at a significant cost if Γ is too large.

We can also observe that the Robust AC-OPF provides a tremendous advantage when compared to the Deterministic AC-OPF, which corresponds to the case with $\Gamma = 0$. For the 14-bus system the average cost savings can be up to 14.4% (with $\Gamma = 1$) and cost standard deviation can be reduced up to 75.2% ($\Gamma = 1.5$), while for the 118-bus system the average cost savings can be up to 17.5% ($\Gamma = 1$) and cost standard deviation can be reduced up to 70.9% ($\Gamma = 1$). Also, for both systems a sufficiently large Γ results in null penalty, while Deterministic AC-OPF presents a significant level of penalty.

In summary: i) a convex relaxation of AC-OPF, such as SOCP-OPF, combined with the alternating direction method provides a good way for approximately solving the difficult max-min problem of finding a worst-case uncertainty realization; ii) the Robust AC-OPF model can be efficiently solved in approximate form; and iii) Robust AC-OPF provides an effective way to manage the uncertainty in active and reactive power demands, as well as in wind and solar power, offering an explicit way to control the desired conservativeness in the operation of the power system.

5.5 Conclusion

We have developed an adaptive robust optimization model for the multi-period AC OPF problem, proposed an effective solution method using convex relaxations of AC OPF, and carried out extensive experiments to understand the benefits and limitations of this approach. The modeling framework developed is innovative in offering a systematic treatment of the non-convexities presented by the power flow equations and the inherent uncertainty in power demand and renewable power, considering technical details such as transmission line capacities and reactive capability curves of

conventional generators and renewable units. The Robust AC-OPF proposed can be efficiently solved and transmission line screening techniques can significantly speed-up the algorithm. Finally, using a realistic simulation platform that mimics the hour to hour operation of the power system, we showed that the Robust AC-OPF can offer significant advantages in terms of cost effectiveness and system reliability, as compared to a Deterministic AC-OPF model, further offering a simple mechanism for balancing the operational conservativeness.

REFERENCES

[1] Amjady, N., Dehghan, S., Attarha, A., and Conejo, A., "Adaptive robust network-constrained AC unit commitment," *IEEE Transactions on Power Systems*, 2016.

[2] Anstreicher, K. M., "Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming," *Journal of Global Optimization*, vol. 43, no. 2-3, pp. 471–484, 2009.

[3] Bai, X., Wei, H., Fujisawa, K., and Wang, Y., "Semidefinite programming for optimal power flow problems," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 6, pp. 383–392, 2008.

[4] Bechert, T. E. and Kwatny, H. G., "On the optimal dynamic dispatch of real power," *IEEE Transactions on Power Apparatus and Systems*, vol. 91, no. 3, pp. 889–898, 1972.

[5] Ben-Tal, A., Boaz, G., and Shimrit, S., "Robust multi-echelon multi-period inventory control," *European Journal of Operational Research*, vol. 199, no. 3, pp. 922–935, 2009.

[6] Ben-Tal, A., El Ghaoui, L., and Nemirovski, A., *Robust optimization*. Princeton University Press, 2009.

[7] Ben-Tal, A., Goryashko, A., Guslitzer, E., and Nemirovski, A., "Adjustable robust solutions of uncertain linear programs," *Mathematical programming, Ser. A*, vol. 99, pp. 351–376, 2004.

[8] Ben-Tal, A. and Nemirovski, A., *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, vol. 2. SIAM, 2001.

[9] Bertsimas, D., Brown, D. B., and Caramanis, C., "Theory and applications of robust optimization," *SIAM Review*, vol. 53, no. 3, pp. 464–501, 2011.

[10] Bertsimas, D., Dunning, I., and Lubin, M., "Reformulations versus cutting planes for robust optimization," *Computational Management Science*, 2015.

[11] Bertsimas, D. and Georghiou, A., "Design of near optimal decision rules in multistage adaptive mixed-integer optimization," *Available at Optimization Online*, 2014.

[12] Bertsimas, D., Iancu, D. A., and Parrilo, P. A., "Optimality of affine policies in multistage robust optimization," *Mathematics of Operations Research*, vol. 35, no. 2, pp. 363–394, 2010.

[13] Bertsimas, D., Litvinov, E., Sun, X. A., Zhao, J., and Zheng, T., "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 52–63, 2013.

[14] Bertsimas, D. and Sim, M., "The price of robustness," *Operations research*, vol. 52, no. 1, pp. 35–53, 2004.

[15] Bienstock, D., Chertkov, M., and Harnett, S., "Chance-constrained optimal power flow: risk-aware network control under uncertainty," *SIAM Review*, vol. 56, no. 3, pp. 461–495, 2014.

[16] Birge, J. R. and Louveaux, F. V., "A multicut algorithm for two-stage stochastic linear programs," *European Journal of Operational Research*, vol. 34, no. 3, pp. 384–392, 1988.

[17] Black, M. and Strbac, G., "Value of bulk energy storage for managing wind power fluctuations," *IEEE Transactions on Energy Conversion*, vol. 22, no. 1, pp. 197–205, 2007.

[18] Blankenship, J. W. and Falk, J. E., "Infinitely constrained optimization problems," *Journal of Optimization Theory and Applications*, vol. 19, no. 2, pp. 261–281, 1976.

[19] Bouffard, F. and Galiana, F. D., "Stochastic security for operations planning with significant wind power generation," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 306–316, 2008.

[20] California ISO, "What are we doing to green the grid?." `http://www.caiso.com/informed/Pages/CleanGrid/default.aspx`, 2016.

[21] Center for energy efficiency and renewable energy, University of Massachusetts Amherst, "RERL wind data." `http://www.ceere.org/rerl/publications/resource\_data/index.html`, 2012.

[22] Chen, X., Sim, M., and Sun, P., "A robust optimization perspective on stochastic programming," *Operations Research*, vol. 55, no. 6, pp. 1058–1071, 2007.

[23] Chen, X., Sim, M., Sun, P., and Zhang, J., "A linear decision-based approximation approach to stochastic programming," *Operations Research*, vol. 56, no. 2, pp. 344–357, 2008.

[24] Chen, X. and Zhang, Y., "Uncertain linear programs: Extended affinely adjustable robust counterparts," *Operations Research*, vol. 57, no. 6, pp. 1469–1482, 2009.

[25] Cheung, K., Gade, D., Silva-Monroy, C., Ryan, S. M., Watson, J.-P., Wets, R. J.-B., and Woodruff, D. L., "Toward scalable stochastic unit commitment: Part 2: solver configuration and performance assessment," *Energy Systems*, 2015.

[26] Conejo, A. J., Carrión, M., and Morales, J. M., *Decision making under uncertainty in electricity markets*, vol. 153. Springer, 2010.

[27] Constantinescu, E. M., Zavala, V. M., Rocklin, M., Lee, S., and Anitescu, M., "A computational framework for uncertainty quantification and stochastic optimization in unit commitment with wind power generation," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 431–441, 2011.

[28] de Oliveira, W. and Sagastizábal, C., "Level bundle methods for oracles with on-demand accuracy," *Optimization Methods and Software*, 2014.

[29] Delage, E. and Iancu, D., "Robust multi-stage decision making," *In INFORMS Tutorials in Operations Research*, pp. 20–46, 2015.

[30] Denholm, P., Jorgenson, J., Hummon, M., Jenkin, T., Palchak, D., Kirby, B., Ma, O., and O'Malley, M., "The value of energy storage for grid applications." `http://www.nrel.gov/docs/fy13osti/58465.pdf`, 2013. NREL Technical Report.

[31] Ela, E. and Kirby, B., "ERCOT event on February 26, 2008 lessons learned," in *Technical Report NREL/TP-500-43373*, National Renewable Energy Laboratory, 2008.

[32] Ellis, A., Nelson, R., VonEngeln, E., Walling, R., McDowell, J., Casey, L., Seymour, E., Peter, W., Barker, C., and Kirby, B., "Reactive power interconnection requirements for pv and wind plants - recommendations to nerc," 2012. SANDIA Technical Report.

[33] FERC, "Recent iso software enhancements and future software and modeling plans." `http://www.ferc.gov/industries/electric/indus-act/rto/rto-iso-soft-2011.pdf`, 2011. FERC Staff Report.

[34] FERC, "Operator-initiated commitment in RTO and ISO markets: Price formation in organized wholesale electricity markets." `http://www.ferc.gov/legal/staff-reports/2014/AD14-14-operator-actions.pdf`, 2014.

[35] Fischetti, M. and Monaci, M., "Cutting plane versus compact formulations for uncertain (integer) linear programs," *Mathematical Programming Computation*, vol. 4, no. 3, pp. 239–273, 2012.

[36] Frank, S. and Rebennack, S., "An introduction to optimal power flow: Theory, formulation, and examples," *IIE Transactions*.

[37] General Electric, "GE Energy 1.5 MW wind turbine," 2009. General Electric Company.

[38] Georghiou, A., Wiesemann, W., and Kuhn, D., "Generalized decision rule approximations for stochastic programming via liftings," *Mathematical Programming*, vol. 152, pp. 301–338, 2015.

[39] Georgia Tech, "ISyE High performance computing facility." http://www.isye.gatech.edu/computers/hpc/, 2012. School of Industrial and Systems Engineering, Georgia Institute of Technology.

[40] Goh, J. and Sim, M., "Robust optimization made easy with rome," *Operations Research*, vol. 59, no. 4, pp. 973–985, 2011.

[41] Gómez-Expósito, A., Conejo, A. J., and Cañizares, C., *Electric energy systems: analysis and operation*. CRC Press, 2008.

[42] Gómez-Expósito, A. and Romero-Ramos, E., "Reliable load flow technique for radial distribution networks," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 1063–1069, 1999.

[43] GWEC, "Global wind statistics 2015." `http://www.gwec.net/wp-content/uploads/vip/GWEC-PRstats-2015_LR_corrected.pdf`, 2016.

[44] Hadjiyiannis, M. J., Goulart, P. J., and Kuhn, D., "An efficient method to estimate the suboptimality of affine controllers," *IEEE Transactions on Automatic Control*, vol. 56, no. 12, pp. 2841–2853, 2011.

[45] Hamilton, J. D., *Time series analysis*, vol. 2. Princeton University Press, 1994.

[46] Heitsch, H. and Römisch, W., "Scenario tree generation for multi-stage stochastic programs," in *Stochastic Optimization Methods in Finance and Energy*, pp. 313–341, Springer, 2011.

[47] Hobbs, B. F., Rothkopf, M. H., O'Neill, R. P., and Chao (Eds.), H., *The Next Generation of Electric Power Unit Commitment Models*, vol. 36 of *International Series in Operations Research & Management Science*. Springer, 2001.

[48] IBM ILOG CPLEX, "V12. 1: User's Manual for CPLEX," *International Business Machines Corporation*, 2009.

[49] Jabr, R. A., "Radial distribution load flow using conic programming," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1458–1459, 2006.

[50] Jabr, R. A., "Optimal power flow using an extended conic quadratic formulation," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1000–1008, 2008.

[51] Jabr, R. A., "Adjustable robust OPF with renewable energy sources," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4742–4751, 2013.

[52] Jabr, R. A., Karaki, S., and Korbane, J. A., "Robust multi-period opf with storage and renewables," *IEEE Transactions on Power Systems*, vol. 30, no. 5, pp. 2790–2799, 2015.

[53] Jiang, R., Wang, J., and Guan, Y., "Robust unit commitment with wind power and pumped storage hydro," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 800–810, 2012.

[54] Jiang, R., Wang, J., Zhang, M., and Guan, Y., "Two-stage minimax regret robust unit commitment," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2271–2282, 2013.

[55] Jiang, R., Zhang, M., Li, G., and Guan, Y., "Two-stage network constrained robust unit commitment problem," *European Journal of Operational Research*, vol. 234, no. 3, pp. 751–762, 2014.

[56] Keshmiri, S. N. and Gao, W., "Multi-objective stochastic economic dispatch," in *North American Power Symposium (NAPS)*, pp. 1–8, 2010.

[57] Keyhani, A., Marwali, M. N., and Dai, M., *Integration of green and renewable energy in electric power systems*. John Wiley & Sons, 2009.

[58] Kocuk, B., Dey, S. S., and Sun, X. A., "Strong SOCP relaxations for the optimal power flow problem," *Operations Research*, 2016.

[59] Kocuk, B., Dey, S. S., and Sun, X. A., "Inexactness of SDP relaxation and valid inequalities for optimal power flow," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 642–651, 2016.

[60] Konno, H., "A cutting plane algorithm for solving bilinear programs," *Mathematical Prgramming*, vol. 11, pp. 14–27, 1976.

[61] Kuhn, D., Wiesemann, W., and Georghiou, A., "Primal and dual linear decision rules in stochastic and robust optimization," *Mathematical Programming*, vol. 130, no. 1, pp. 177–209, 2011.

[62] Lavaei, J. and Low, S. H., "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, 2012.

[63] Lee, C., Liu, C., Mehrotra, S., and Shahidehpour, M., "Modeling transmission line constraints in two-stage robust unit commitment problem," *IEEE Transactions on Power Systems*, 2013.

[64] Lee, Y. and Baldick, R., "A frequency-constrained stochastic economic dispatch model," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2301–2312, 2013.

[65] Li, Y., Li, W., Yan, W., Yu, J., and Zhao, X., "Probabilistic optimal power flow considering correlations of wind speeds following different distributions," *IEEE Transactions on Power Systems*, vol. 29, no. 4, pp. 1847–1854, 2014.

[66] Li, Z., Wu, W., Zhang, B., and Wang, B., "Adjustable robust real-time power dispatch with large-scale wind power integration," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 2, pp. 357–368, 2015.

[67] Lorca, A. and Sun, X. A., "Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1702–1713, 2015.

[68] Lorca, A. and Sun, X. A., "Multistage robust unit commitment with dynamic uncertainty sets and energy storage," *Available at Optimization Online*, 2016.

[69] Lorca, A., Sun, X. A., Litvinov, E., and Zheng, T., "Multistage adaptive robust optimization for the unit commitment problem," *Operations Research*, vol. 64, no. 1, pp. 32–51, 2016.

[70] Low, S. H., "Convex relaxation of optimal power flow, part i: Formulations and equivalence," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 15–27, 2014.

[71] Lubin, M., Dvorkin, Y., and Backhaus, S., "A robust approach to chance constrained optimal power flow with renewable generation," *IEEE Transactions on Power Systems*, 2015.

[72] Minoux, M., "Two-stage robust optimization, state-space representable uncertainty and applications," *RAIRO-Operations Research*, vol. 48, no. 4, pp. 455–475, 2014.

[73] Molzahn, D. K. and Hiskens, I., "Sparsity-exploiting moment-based relaxations of the optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3168–3180, 2015.

[74] Morales, J. M., Conejo, A. J., and Pérez-Ruiz, J., "Economic valuation of reserves in power systems with high penetration of wind power," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 900–910, 2009.

[75] Morales, J. M., Minguez, R., and Conejo, A. J., "A methodology to generate statistically dependent wind speed scenarios," *Applied Energy*, vol. 87, no. 3, pp. 843–855, 2010.

[76] Navid, N. and Rosenwald, G., "Market solutions for managing ramp flexibility with high penetration of renewable resource," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 4, pp. 784–790, 2012.

[77] NREL, "Transmission grid integration — western wind dataset," 2012. http://www.nrel.gov/electricity/transmission/wind_integration_dataset.html.

[78] Ortega-Vazquez, M. A. and Kirschen, D. S., "Estimating the spinning reserve requirements in systems with significant wind power generation penetration," *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 114–123, 2009.

[79] Ostrowski, J., Anjos, M. F., and Vannelli, A., "Tight mixed integer linear programming formulations for the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, p. 39, 2012.

[80] Ozturk, U. A., Mazumdmar, M., and Norman, B. A., "A solution to the stochastic unit commitment problem using chance constrained programming," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1589–1598, 2004.

[81] Papavasiliou, A., Oren, S., and O'Neill, R. P., "Reserve requirements for wind power integration: A scenario-based stochastic programming framework," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2197–2206, 2011.

[82] Papavasiliou, A. and Oren, S. S., "Multi-area stochastic unit commitment for high wind penetration in a transmission constrained network," *Operations Research*, vol. 61, no. 3, pp. 578–592, 2013.

[83] Papavasiliou, A., Oren, S. S., and Rountree, B., "Applying high performance computing to transmission-constrained stochastic unit commitment for renewable energy integration," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1109–1120, 2015.

[84] Phan, D. and Ghosh, S., "Two-stage stochastic optimization for optimal power flow under renewable generation uncertainty," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 24, no. 1, p. 2, 2014.

[85] Potter, C. W., Lew, D., McCaa, J., Cheng, S., Eichelberger, S., and Grimit, E., "Creating the dataset for the western wind and solar integration study (USA)," *Wind Engineering*, vol. 32, no. 4, pp. 325–338, 2008.

[86] Reinsel, G. C., *Elements of multivariate time series analysis*. Springer Science & Business Media, 2003.

[87] Ross, D. W. and Kim, S., "Dynamic economic dispatch of generation," *IEEE Transactions on Power Apparatus and Systems*, vol. 99, no. 6, pp. 2060–2068, 1980.

[88] Ruiz, P. A., Philbrick, C. R., and Sauer, P. W., "Wind power day-ahead uncertainty management through stochastic unit commitment policies," *IEEE Power Systems Conference and Exposition (PSCE'09)*, 2009.

[89] Ruiz, P. A., Philbrick, C. R., Zak, E., Cheung, K. W., and Sauer, P. W., "Uncertainty management in the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 642–651, 2009.

[90] Sahinidis, N. V., "Baron: A general purpose global optimization software package," *Journal of global optimization*, vol. 8, no. 2, pp. 201–205, 1996.

[91] Shapiro, A., Dentcheva, D., and Ruszczyński, A. P., *Lectures on stochastic programming: modeling and theory*, vol. 9. SIAM, 2009.

[92] Sharifzadeh, H. and Amjady, N., "Stochastic security-constrained optimal power flow incorporating preventive and corrective actions," *International Transactions on Electrical Energy Systems*, 2016.

[93] Solar Power Europe, "Global wind statistics 2015." `http://helapco.gr/pdf/Global_Market_Outlook_2015_-2019_lr_v23.pdf`, 2015.

[94] Street, A., Oliveira, F., and Arroyo, J. M., "Contingency-constrained unit commitment with $n - k$ security criterion: A robust optimization approach," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1581–1590, 2011.

[95] Summers, T., Warrington, J., Morari, M., and Lygeros, J., "Stochastic optimal power flow based on conditional value at risk and distributional robustness," *International Journal of Electrical Power & Energy Systems*, vol. 72, pp. 116–125, 2015.

[96] Sun, X. A. and Lorca, A., "Adaptive robust optimization for daily power system operation," in *Power Systems Computation Conference (PSCC)*, pp. 1–9, IEEE, 2014.

[97] Takriti, S., Birge, J. R., and Long, E., "A stochastic model for the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 11, no. 3, pp. 1497–1508, 1996.

[98] Tuohy, A., Meibom, P., Denny, E., and O'Malley, M., "Unit commitment for systems with significant wind penetration," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 592–601, 2009.

[99] Tuy, H., *Convex analysis and global optimization*. Kluwer Academic Publishers, 1998.

[100] Wang, H., Murillo-Sánchez, C. E., Zimmerman, R. D., and Thomas, R. J., "On computational issues of market based optimal power flow," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1185–1193, 2007.

[101] Wang, J., Shahidehpour, M., and Li, Z., "Security-constrained unit commitment with volatile wind power generation," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1319–1327, 2008.

[102] Wang, Q., Guan, Y., and Wang, J., "A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 206–215, 2012.

[103] Wang, Q., Watson, J. P., and Guan, Y., "Two-stage robust optimization for $n - k$ contingency-constrained unit commitment," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2366–2375, 2013.

[104] Warrington, J., Goulart, P. J., Mariéthoz, S., and Morari, M., "Robust reserve operation in power systems using affine policies," in *51st IEEE Conference on Decision and Control*, pp. 1111–1117, IEEE, 2012.

[105] Warrington, J., Goulart, P. J., Mariéthoz, S., and Morari, M., "Policy-based reserves for power systems," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4427–4437, 2013.

[106] Warrington, J., Hohl, C., Goulart, P. J., and Morari, M., "Optimal unit commitment accounting for robust affine reserve policies," in *2014 American Control Conference*, pp. 5049–5055, ACC, 2014.

[107] Warrington, J., Hohl, C., Goulart, P. J., and Morari, M., "Rolling unit commitment and dispatch with multi-stage recourse policies for heterogeneous devices," *IEEE Transactions on Power Systems*, 2015.

[108] Wu, L., Shahidehpour, M., and Li, T., "Stochastic security-constrained unit commitment," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 800–811, 2007.

[109] Wu, W., Chen, J., Zhang, B., and Sun, H., "A robust wind power optimization method for look-ahead power dispatch," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 2, pp. 507–515, 2014.

[110] Xie, L., Carvalho, P., Ferreira, L., Liu, J., Krogh, B., Popli, N., and Ilic, M., "Wind integration in power systems: Operational challenges and possible solutions," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 214–232, 2011.

[111] Xie, L., Gu, Y., X., Z., and Genton, M., "Short-term spatio-temporal wind power forecast in robust look-ahead power system dispatch," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 511–520, 2013.

[112] Xie, L., Gu, Y., Zhu, X., and Genton, M., "Power system economic dispatch with spatio-temporal wind forecasts," in *Energytech, 2011 IEEE*, pp. 1–6, IEEE, 2011.

[113] Zeng, B. and Zhao, L., "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Operations Research Letters*, vol. 41, no. 5, pp. 457–461, 2013.

[114] Zhang, Y., Gatsis, N., and Giannakis, G. B., "Risk-constrained energy management with multiple wind farms," in *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*, pp. 1–6, IEEE, 2013.

[115] Zhao, C. and Guan, Y., "Unified stochastic and robust unit commitment," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3353–3361, 2013.

[116] Zhao, C., Wang, J., Watson, J. P., and Guan, Y., "Multi-stage robust unit commitment considering wind and demand response uncertainties," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2708–2717, 2013.

[117] Zhao, J., Zheng, T., and Litvinov, E., "Variable resource dispatch through do-not-exceed limit," *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 820–828, 2015.

[118] Zhao, L. and Zeng, B., "Robust unit commitment problem with demand response and wind energy," in *IEEE Power and Energy Society General Meeting*, 2012.

[119] Zheng, T., Zhao, J., Litvinov, E., and Zhao, F., "Robust optimization and its application to power system operation," in *CIGRE*, 2012.

[120] Zimmerman, R. D., Murillo-Sánchez, C. E., and Thomas, R. J., "MATPOWER: Steady-state operation, planning and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

VITA

Álvaro Lorca was born in Santiago, Chile, in 1985. In 2004, he became a student at Pontificia Universidad Católica de Chile, where he later received the Industrial Engineering degree with specialization in Mathematical engineering, and the M.Sc. degree in Industrial Engineering. In 2011, Álvaro moved to Atlanta to pursue doctoral studies in Operations Research at the Georgia Institute of Technology. Under the supervision of Dr. Xu Andy Sun, he worked on Robust Optimization models and algorithms for electric power system operations, and also collaborated with ISO New England on these topics. Beyond his work in energy, he has also worked with Drs. Pinar Keskinocak and Ozlem Ergun in post-disaster logistical operations. During his doctoral studies, Álvaro was awarded third place in the INFORMS JFIG paper competition in 2014, and honorable mention of the Alice and John Harvis Ph.D. student research award in 2016, awarded by the Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology.

On a personal level, Álvaro is married and has one daughter.